

# (12) UK Patent Application (19) GB (11) 2 357 767 (13) A

(43) Date of A Publication 04.07.2001

(21) Application No 0023315.5

(22) Date of Filing 22.09.2000

## (30) Priority Data

(31) 60155493	(32) 23.09.1999	(33) US
(31) 09416901	(32) 13.10.1999	
(31) 09404133	(32) 23.09.1999	
(31) 60169232	(32) 06.12.1999	
(31) 99120881	(32) 23.09.1999	(33) WO

## (71) Applicant(s)

Pharmacia & Upjohn Company  
(Incorporated in USA - Delaware)  
301 Henrietta Street, Kalamazoo, Michigan 49001,  
United States of America

## (72) Inventor(s)

Michael Jerome Bienkowski  
Mark Gurney

(51) INT CL<sup>7</sup>

C07K 14/47 , A61K 38/00 , A61P 25/28 , C12N 1/21  
9/64 15/57 , C12Q 1/68 , G01N 33/68

## (52) UK CL (Edition S )

C3H HA5 H339  
G1B BAB B734  
U1S S2417

## (56) Documents Cited

EP 0848062 A DE 019910108 A DE 019849073 A  
J. NEUROCHEM, 1998, 71, 1920-1925, GK GOURAS  
ET AL

## (58) Field of Search

ONLINE: EPODOC, WPI, JAPIO, DGENE, CAPLUS,  
BIOSIS, MEDLINE

## (74) Agent and/or Address for Service

Gill Jennings & Every  
Broadgate House, 7 Eldon Street, LONDON,  
EC2M 7LH, United Kingdom

## (54) Abstract Title

**Methods of assaying human Aspartyl protease 1 alpha-secretase activity**

(57) A method for assaying hu-Asp1  $\alpha$ -secretase activity is claimed, comprising contacting hu-Asp1 protein with an amyloid precursor protein (APP) substrate containing an  $\alpha$ -secretase cleavage site and measuring the cleavage of said substrate at said site. The hu-Asp1 amino acid sequence may comprise the sequence of SEQ ID NO:2 or preferably modified sequences are claimed lacking the amino terminal domain, transmembrane domain or cytoplasmic domain. Methods of identifying modulators of hu-Asp1  $\alpha$ -secretase activity as candidate Alzheimer's disease modulators and their use in therapeutic compositions are also claimed. Preferably the APP substrate contains the  $\alpha$ -secretase cleavage site LVFFAEDF or KLVFFAED. Alternatively claimed is a hu-Asp1 substrate peptide comprising 50 or fewer amino acids including the cleavage site GLALALEP and methods of assaying hu-Asp1 proteolytic cleavage. The Hu-Asp2  $\beta$ -secretase gene, polypeptide and assay for  $\beta$ -secretase activity is also described.

GB 2 357 767 A

FIGURE 1 (1)

ATGGGCGCACTGGCCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC  
 M G A L A R A L L L P L L A Q W L L R A  
 CCCCCGAGCTGGCCCCCGCGCCCTTACGCTGCCCCCTCCGGGTGGCCGCGGCCACGAAC  
 A P E L A P A P F T L P L R V A A A T N  
 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG  
 R V V A P T P G P G T P A E R H A D G L  
 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGGGGCGCCGCAACTTCTTGGCCATG  
 A L A L E P A L A S P A G A A N F L A M  
 GTAGACAACCTGCAGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC  
 V D N L Q G D S G R G Y Y L E M L I G T  
 CCCCCGAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACCTTGCCGTGGCAGGA  
 P P Q K L Q I L V D T G S S N F A V A G  
 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC  
 T P H S Y I D T Y F D T E R S S T Y R S  
 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA  
 K G F D V T V K Y T Q G S W T G F V G E  
 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTC AACATTGCCACTATT  
 D L V T I P K G F N T S F L V N I A T I  
 TTTGAATCAGAGAATTTCTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT  
 F E S E N F F L P G I K W N G I L G L A  
 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA  
 Y A T L A K P S S S L E T F F D S L V T  
 CAAGCAAACATCCCCAACGTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT  
 Q A N I P N V F S M Q M C G A G L P V A  
 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTGTATAAA  
 G S G T N G G S L V L G G I E P S L Y K  
 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATCTGAAA  
 G D I W Y T P I K E E W Y Y Q I E I L K  
 TTGGAAATTTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC  
 L E I G G Q S L N L D C R E Y N A D K A  
 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG  
 I V D S G T T L L R L P Q K V F D A V V  
 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC  
 E A V A R A S L I P E F S D G F W T G S  
 CAGCTGGCGTGCTGGACGAATTCGGAACACCTTGGTCTTACTTCCCTAAAATCTCCATC  
 Q L A C W T N S E T P W S Y F P K I S I  
 TACCTGAGAGATGAGAACTCCAGCAGGTCAATCCGTATCACAATCCTGCCTCAGCTTTAC  
 Y L R D E N S S R S F R I T I L P Q L Y  
 ATTCAGCCCATGATGGGGGCGGCCTGAATTATGAATGTTACCGATTCCGGCATTTCGCCA  
 I Q P M M G A G L N Y E C Y R F G I S P  
 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC  
 S T N A L V I G A T V M E G F Y V I F D  
 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA



FIGURE 1 (2)

R A Q K R V G F A A S P C A E I A G A A  
GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT  
V S E I S G P F S T E D V A S N C V P A  
CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA  
Q S L S E P I L W I V S Y A L M S V C G  
GCCATCCTCCTTGTCCTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTGAGCGTCGCCCC  
A I L L V L I V L L L L P F R C Q R R P  
CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAATGAATA  
R D P E V V N D E S S L V R H R W K  
GCCAGGCCTGACCTCAAGCAACCATGAACTCAGCTATTAAGAAAATCACATTTCCAGGGC  
AGCAGCCGGGATCGATGGTGGCGCTTTCTCCTGTGCCCACCCGCTTCAATCTCTGTTCT  
GCTCCCAGATGCCCTTCTAGATTCACTGTCTTTTGATTCTTGATTTCAGCTTTCAAATC  
CTCCCTACTTCCAAGAAAAATAATTAAAAAAAACCTTCATTCTAAACCAAAAAAAAAA  
AAAA



FIGURE 2 (1)

ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAGTGCTGCCTGCCAC  
 M A Q A L P W L L L W M G A G V L P A H  
 GGCACCCAGCACGGCATCCGGCTGCCCCTGCGCAGCGGCCTGGGGGGCGCCCCCTGGGG  
 G T Q H G I R L P L R S G L G G A P L G  
 CTGCGGCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCCGGCCGGAGGGGCAGCTTT  
 L R L P R E T D E E P E E P G R R G S F  
 GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC  
 V E M V D N L R G K S G Q G Y Y V E M T  
 GTGGGCAGCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACTTTGCA  
 V G S P P Q T L N I L V D T G S S N F A  
 GTGGGTGCTGCCCCCCCACCCCTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA  
 V G A A P H P F L H R Y Y Q R Q L S S T  
 TACCGGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG  
 Y R D L R K G V Y V P Y T Q G K W E G E  
 CTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATT  
 L G T D L V S I P H G P N V T V R A N I  
 GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAAGTGGGAAGGCATCCTG  
 A A I T E S D K F F I N G S N W E G I L  
 GGGCTGGCCTATGCTGAGATTGCCAGGCTTTGTGGTGCTGGCTTCCCCCTCAACCAAGTCT  
 G L A Y A E I A R L C G A G F P L N Q S  
 GAAGTGCTGGCCTCTGTGCGAGGGAGCATGATCATTGGAGGTATCGACCACTCGCTGTAC  
 E V L A S V G G S M I I G G I D H S L Y  
 ACAGGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTATGAGGTGATCATTGTG  
 T G S L W Y T P I R R E W Y Y E V I I V  
 CGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAGTACAACATATGACAAG  
 R V E I N G Q D L K M D C K E Y N Y D K  
 AGCATTTGTGGACAGTGGCACCACCAACCTTCGTTTGGCCAAGAAAGTGTGTTGAAGCTGCA  
 S I V D S G T T N L R L P K K V F E A A  
 GTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGA  
 V K S I K A A S S T E K F P D G F W L G  
 GAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATTTTCCCAGTCATCTCA  
 E Q L V C W Q A G T T P W N I F P V I S  
 CTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTCCGCATCACCATCCTTCCGCAGCAA  
 L Y L M G E V T N Q S F R I T I L P Q Q  
 TACCTGCGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGTTACAAGTTTGCCATC

FIGURE 2 (2)

Y L R P V E D V A T S Q D D C Y K F A I  
TCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTC  
S Q S S T G T V M G A V I M E G F Y V V  
TTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGCCATGTGCACGATGAG  
F D R A R K R I G F A V S A C H V H D E  
TTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTACCTTGGACATGGAAGACTGTGGCTAC  
F R T A A V E G P F V T L D M E D C G Y  
AACATTCCACAGACAGATGAGTCAACCCTCATGACCATAGCCTATGTCATGGCTGCCATC  
N I P Q T D E S T L M T I A Y V M A A I  
TGCGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGTGTCAGTGGCGCTGCCTCCGCTGC  
C A L F M L P L C L M V C Q W R C L R C  
CTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTGAAGTGAGGAGGCCCA  
L R Q Q H D D F A D D I S L L K  
TGGGCAGAAGATAGAGATTCCCCTGGACCACACCTCCGTGGTTCACCTTTGGTCACAAGTA  
GGAGACACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCTCCCCACCCACCAAATGC  
CTCTGCCTTGATGGAGAAGGAAAAGGCTGGCAAGGTGGGTTCAGGGACTGTACCTGTAG  
GAAACAGAAAAGAGAAGAAAGAAGCACTCTGCTGGCGGGAATACTCTTGGTCACCTCAA  
TTTAAGTCGGGAAATTCTGCTGCTTGAAACTTCAGCCCTGAACCTTTGTCCACCATTCCT  
TTAAATTCTCCAACCCAAAGTATTCTTCTTTTCTTAGTTTCAGAAGTACTGGCATCACAC  
GCAGGTTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAAGAGACCAAGCTTGTTT  
CCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGCTATTTGCTTTAGAGACAGG  
GACTGTATAAACAAGCCTAACATTGGTGCAAAGATTGCCTCTTGAAAAAAAAAAAAA

FIGURE 3 (1)

ATGGCCCAAGCCCTGCCCTGGCTCCTGCTGTGGATGGGCGCGGGAGTGCTGCCTGCCAC  
 M A Q A L P W L L L W M G A G V L P A H  
 GGCACCCAGCACGGCATCCGGCTGCCCCTGCGCAGCGGCCTGGGGGGCGCCCCCTGGGG  
 G T Q H G I R L P L R S G L G G A P L G  
 CTGCGGCTGCCCCGGGAGACCGACGAAGAGCCCCGAGGAGCCCGGCCGGAGGGGCAGCTTT  
 L R L P R E T D E E P E E P G R R G S F  
 GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC  
 V E M V D N L R G K S G Q G Y Y V E M T  
 GTGGGCAGCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACCTTTGCA  
 V G S P P Q T L N I L V D T G S S N F A  
 GTGGGTGCTGCCCCCACCCTTCTGTCATCGCTACTACCAGAGGCAGCTGTCCAGCACA  
 V G A A P H P F L H R Y Y Q R Q L S S T  
 TACCGGGACCTCCGGAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG  
 Y R D L R K G V Y V P Y T Q G K W E G E  
 CTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATT  
 L G T D L V S I P H G P N V T V R A N I  
 GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGGAAGGCATCCTG  
 A A I T E S D K F F I N G S N W E G I L  
 GGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT  
 G L A Y A E I A R P D D S L E P F F D S  
 CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGGCTTC  
 L V K Q T H V P N L F S L Q L C G A G F  
 CCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTCCGAGGGAGCATGATCATTTGGAGGTATC  
 P L N Q S E V L A S V G G S M I I G G I  
 GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT  
 D H S L Y T G S L W Y T P I R R E W Y Y  
 GAGGTCATCATTTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG  
 E V I I V R V E I N G Q D L K M D C K E  
 TACAACTATGACAAGAGCATTGTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAA  
 Y N Y D K S I V D S G T T N L R L P K K  
 GTGTTTGAAGCTGCAGTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGAT  
 V F E A A V K S I K A A S S T E K F P D

FIGURE 3 (2)

GGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATT  
G F W L G E Q L V C W Q A G T T P W N I  
TTCCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTCCGCATCACC  
F P V I S L Y L M G E V T N Q S F R I T  
ATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGCCACGTCCCAAGACGACTGT  
I L P Q Q Y L R P V E D V A T S Q D D C  
TACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG  
Y K F A I S Q S S T G T V M G A V I M E  
GGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGC  
G F Y V V F D R A R K R I G F A V S A C  
CATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAGGCCCTTTTGTACCTTGGACATG  
H V H D E F R T A A V E G P F V T L D M  
GAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCAACCCTCATGACCATAGCCTAT  
E D C G Y N I P Q T D E S T L M T I A Y  
GTCATGGCTGCCATCTGCGCCCTCTTCATGCTGCCACTCTGCCTCATGGTGTGTCAGTGG  
V M A A I C A L F M L P L C L M V C Q W  
CGCTGCCTCCGCTGCCTGCGCCAGCAGCATGATGACTTTGCTGATGACATCTCCCTGCTG  
R C L R C L R Q Q H D D F A D D I S L L  
AAGTGAGGAGGCCCATGGGCAGAAGATAGAGATTCCCCTGGACCACACCTCCGTGGTTCA  
K  
CTTTGGTCACAAGTAGGAGACACAGATGGCACCTGTGGCCAGAGCACCTCAGGACCCTCC  
CCACCCACCAAATGCCTCTGCCTTGATGGAGAAGGAAAAGGCTGGCAAGGTGGGTTCAG  
GGACTGTACCTGTAGGAAACAGAAAAGAGAAGAAAGAACACTCTGCTGGCGGGAATACT  
CTTGGTCACCTCAAATTTAAGTCGGGAAATTCCTGCTGCTTGAAACTTCAGCCCTGAACCT  
TTGTCCACCATTCCTTTAAATTCCTCAACCCAAAGTATTCTTCTTTTCTTAGTTTCAGAA  
GTACTGGCATCACACGCAGGTTACCTTGGCGTGTGTCCCTGTGGTACCCTGGCAGAGAAG  
AGACCAAGCTTGTTTCCCTGCTGGCCAAAGTCAGTAGGAGAGGATGCACAGTTTGCTATT  
TGCTTTAGAGACAGGGACTGTATAACAAGCCTAACATTGGTGCAAAGATTGCCTCTTGA  
ATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 4

ATGGCCCCAGCGCTGCACTGGCTCCTGCTATGGGTGGGCTCGGGAATGCTGCCTGCCCAG  
 M A P A L H W L L L W V G S G M L P A Q  
 GGAACCCATCTCGGCATCCGGCTGCCCTTCGCAGCGGCTGGCAGGGCCACCCCTGGGC  
 G T H L G I R L P L R S G L A G P P L G  
 CTGAGGCTGCCCCGGGAGACTGACGAGGAATCGGAGGAGCCTGGCCGGAGAGGCAGCTTT  
 L R L P R E T D E E S E E P G R R G S F  
 GTGGAGATGGTGGACAACCTGAGGGGAAAGTCCGGCCAGGGCTACTATGTGGAGATGACC  
 V E M V D N L R G K S G Q G Y Y V E M T  
 GTAGGCAGCCCCCACAGACGCTCAACATCCTGGTGGACACGGGCAGTAGTAACCTTTGCA  
 V G S P P Q T L N I L V D T G S S N F A  
 GTGGGGGTGCCCCACACCCTTTCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA  
 V G A A P H P F L H R Y Y Q R Q L S S T  
 TATCGAGACCTCCGAAAGGGTGTGTATGTGCCCTACACCCAGGGCAAGTGGGAGGGGGAA  
 Y R D L R K G V Y V P Y T Q G K W E G E  
 CTGGGCACCGACCTGGTGAGCATCCCTCATGGCCCCAACGTCACTGTGCGTGCCAACATT  
 L G T D L V S I P H G P N V T V R A N I  
 GCTGCCACTGAATCGGACAAGTTCTTCATCAATGGTTCCAACCTGGGAGGGGCATCCTA  
 A A I T E S D K F F I N G S N W E G I L  
 GGGCTGGCCTATGCTGAGATTGCCAGGCCGACGACTCTTTGGAGCCCTTCTTTGACTCC  
 G L A Y A E I A R P D D S L E P F F D S  
 CTGGTGAAGCAGACCCACATTCCCAACATCTTTTCCCTGCAGCTCTGTGGCGCTGGCTTC  
 L V K Q T H I P N I F S L Q L C G A G F  
 CCCCTCAACCAGACCGAGGCACTGGCCTCGGTGGGAGGGAGCATGATCATTGGTGGTATC  
 P L N Q T E A L A S V G G S M I I G G I  
 GACCACTCGTATACACGGGCAGTCTCTGGTACACACCCATCCGGCGGGAGTGGTATTAT  
 D H S L Y T G S L W Y T P I R R E W Y Y  
 GAAGTGATCATTGTACGTGTGGAATCAATGGTCAAGATCTCAAGATGGACTGCAAGGAG  
 E V I I V R V E I N G Q D L K M D C K E  
 TACAACACGACAAGAGCATTGTGGACAGTGGGACCACCAACCTTCGCTTGCCCAAGAAA  
 Y N Y D K S I V D S G T T N L R L P K K  
 GTATTTGAAGCTGCCGTCAAGTCCATCAAGGCAGCCTCCTCGACGGAGAAGTTCCCGGAT  
 V F E A A V K S I K A A S S T E K F P D  
 GGCTTTTGGCTAGGGGAGCAGCTGGTGTGCTGGCAAGCAGGCACGACCCCTTGGAAACATT  
 G F W L G E I L V C W Q A G T T P W N I  
 TTCCAGTCATTTCACCTTTACCTCATGGGTGAAGTCACCAATCAGTCTTCCGCATCACC  
 F P V I S L Y L M G E V T N Q S F R I T  
 ATCCTTCCTCAGCAATACCTACGGCCGGTGGAGGACGTGGCCACGTCCCAAGACGACTGT  
 I L P Q Q Y L R P V E D V A T S Q D D C  
 TACAAGTTGCTGTCTCACAGTCATCCACGGGCACTGTTATGGGAGCCGTCATCATGGAA  
 Y K F A V S Q S S T G T V M G A V I M E  
 GGTTTCTATGTCGCTTTCGATCGAGCCCCGAAAGCGAATTGGCTTTGCTGTGACGCTTGC  
 G F Y V V F D R A R K R I G F A V S A C  
 CATGTGCACGATGAGTTCAGGACGGCGGCACTGGAAGGTCCGTTTGTACGGCAGACATG  
 H V H D E F R T A A V E G P F V T A D M  
 GAAGACTGTGGCTACAACATTCCCCAGACAGATGAGTCAACACTTATGACCATAGCCTAT  
 E D C G Y N I P Q T D E S T L M T I A Y  
 GTCATGGCGGCCATCTGCGCCCTCTTCATGTTGCCACTCTGCCTCATGGTATGTCAGTGG  
 V M A A I C A L F M L P L C L M V C Q W  
 CGCTGCCTGCGTTGCTGCGCCACCAGCAGATGACTTTGCTGATGACATCTCCCTGCTC  
 R C L R C L R H Q H D D F A D D I S L L  
 AAGTAAGGAGGCTCGTGGGCAGATGATGGAGACGCCCCCTGGACCACATCTGGGTGGTTCC  
 K  
 CTTTGGTTCACATGAGTTGGAGCTATGGATGGTACCTGTGGCCAGAGCACCTCAGGACCCT  
 CACCAACCTGCCAATGCTTCTGGCGTGACAGAACAGAGAAATCAGGCAAGCTGGATTACA  
 GGGCTTGACCTGTAGGACACAGGAGAGGGAAGGAAGCAGCGTTCTGGTGGCAGGAATAT  
 CTTAGGCACCACAACTTGAGTTGGAAATTTTGTGCTTGAAGCTTCAGCCCTGACCCT  
 CTGCCCAGCATCCTTTAGAGTCTCCAACCTAAAGTATTCTTTATGTCTTCCAGAAGTAC  
 TGGCGTCATACTCAGGCTACCCGGCATGTGTCCCTGTGGTACCCTGGCAGAGAAAGGGCC  
 AATCTCATTCCCTGCTGGCCAAAGTCAGCAGAAGAAGGTGAAGTTTGCCAGTTGCTTTAG  
 TGATAGGAGCTGCAGACTCAAGCCTACACTGGTACAAAGACTGCGTCTTGAGATAAAACAA  
 GAA

FIGURE 5

```

1 MAQALPWLLLWMGAGVLPAGHTQHGIRLPLRSGLGAPLGLRLPRETDEE 50
  |||||
1 MAPALHWLLLWVGSGMLPAQGTHLGIRLPLRSGLAGPPLGLRLPRETDEE 50
  |||||

51 PEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFA 100
  |||||
51 SEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFA 100
  |||||

101 VGAAPHPFLHRYYQRLSSTYRDLRGVYVPYTQGWEGELGTDLVSIPH 150
  |||||
101 VGAAPHPFLHRYYQRLSSTYRDLRGVYVPYTQGWEGELGTDLVSIPH 150
  |||||

151 GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS 200
  |||||
151 GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS 200
  |||||

201 LVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLW 250
  |||||:|||||
201 LVKQTHIPNIFSLQLCGAGFPLNQTEALASVGGSMIIGGIDHSLYTGSLW 250
  |||||:|||||

251 YTPIRREWYVEVIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPLPKK 300
  |||||
251 YTPIRREWYVEVIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPLPKK 300
  |||||

301 VFEEAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG 350
  |||||
301 VFEEAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG 350
  |||||

351 EVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME 400
  |||||:|||||
351 EVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIME 400
  |||||:|||||

401 GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQT 450
  |||||
401 GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQT 450
  |||||

451 DESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLRQHQHDDFADDISLL 500
  |||||
451 DESTLMTIAYVMAAICALFMLPLCLMVCQWRCLRCLRHQHQHDDFADDISLL 500
  |||||

501 K 501
  |
501 K 501

```

FIGURE 6 (1)

ATGGCTAGCATGACTGGTGGACAGCAAATGGGTGCGGGATCCACCCAGCACGGCATCCGG  
 M A S M T G G Q Q M G R G S T Q H G I R  
 CTGCCCCCTGCGCAGCGGCCTGGGGGGCGCCCCCTGGGGCTGCGGCTGCCCCGGGAGACC  
 L P L R S G L G G A P L G L R L P R E T  
 GACGAAGAGCCCCGAGGAGCCCGCGGAGGGGAGCTTTGTGGAGATGGTGGACAACCTG  
 D E E P E E P G R R G S F V E M V D N L  
 AGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACCGTGGGCAGCCCCCGCAGACG  
 R G K S G Q G Y Y V E M T V G S P P Q T  
 CTCAACATCCTGGTGGATACAGGCAGCAGTAACCTTGCAGTGGGTGCTGCCCCCACCCC  
 L N I L V D T G S S N F A V G A A P H P  
 TTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACATACCGGGACCTCCGGAAGGGC  
 F L H R Y Y Q R Q L S S T Y R D L R K G  
 GTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGC  
 V Y V P Y T Q G K W E G E L G T D L V S  
 ATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAACATTGCTGCCATCACTGAATCAGAC  
 I P H G P N V T V R A N I A A I T E S D  
 AAGTTCTTCATCAACGGCTCCAAGTGGGAAGGCATCCTGGGGCTGGCCTATGCTGAGATT  
 K F F I N G S N W E G I L G L A Y A E I  
 GCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACGTT  
 A R P D D S L E P F F D S L V K Q T H V  
 CCCAACCTCTTCTCCCTGCAGCTTTGTGGTGCTGGCTTCCCCCTCAACCAGTCTGAAGTG  
 P N L F S L Q L C G A G F P L N Q S E V  
 CTGGCCTCTGTGCGGAGGGAGCATGATCATTGGAGGTATCGACCACTCGCTGTACACAGGC  
 L A S V G G S M I I G G I D H S L Y T G  
 AGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTATGAGGTCACTATTGTGCGGGTG  
 S L W Y T P I R R E W Y Y E V I I V R V  
 GAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAGTACAACCTATGACAAGAGCATT  
 E I N G Q D L K M D C K E Y N Y D K S I  
 GTGGACAGTGGCACCACCAACCTTCGTTTGCCCCAAGAAAGTGTGGAAGCTGCAGTCAAA  
 V D S G T T N L R L P K K V F E A A V K  
 TCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGAGAGCAG  
 S I K A A S S T E K F P D G F W L G E Q  
 CTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAAACATTTTCCAGTCATCTCACTCTAC  
 L V C W Q A G T T P W N I F P V I S L Y  
 CTAATGGGTGAGGTTACCAACCAGTCCTTCCGCATCACCATCCTTCCGCAGCAATACCTG  
 L M G E V T N Q S F R I T I L P Q Q Y L  
 CGGCCAGTGAAGATGTGGCCACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAG

FIGURE 6 (2)

R P V E D V A T S Q D D C Y K F A I S Q  
TCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGAT  
S S T G T V M G A V I M E G F Y V V F D  
CGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGCCATGTGCACGATGAGTTCAGG  
R A R K R I G F A V S A C H V H D E F R  
ACGGCAGCGGTGGAAGGCCCTTTTGTACCTTGGACATGGAAGACTGTGGCTACAACATT  
T A A V E G P F V T L D M E D C G Y N I  
CCACAGACAGATGAGTCATGA  
P Q T D E S \*



FIGURE 7 (1)

ATGGCTAGCATGACTGGTGGACAGCAAATGGGTGCGGATCGATGACTATCTCTGACTCT  
 M A S M T G G Q Q M G R G S M T I S D S  
 CCGCGTGAACAGGACGGATCCACCCAGCACGGCATCCGGCTGCCCCGCGCAGCGGCCTG  
 P R E Q D G S T Q H G I R L P L R S G L  
 GGGGCGCCCCCTGGGGCTGCGGCTGCCCCGGGAGACCGACGAAGAGCCCGAGGAGCCC  
 G G A P L G L R L P R E T D E E P E E P  
 GGCCGGAGGGGCGAGCTTTGTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGC  
 G R R G S F V E M V D N L R G K S G Q G  
 TACTACGTGGAGATGACCGTGGGCAGCCCCCGCAGACGCTCAACATCCTGGTGGATACA  
 Y Y V E M T V G S P P Q T L N I L V D T  
 GGCAGCAGTAACTTTGCAGTGGGTGCTGCCCCCACCCTTCCTGCATCGCTACTACCAG  
 G S S N F A V G A A P H P F L H R Y Y Q  
 AGGCAGCTGTCCAGCACATACCGGGACCTCCGGAAGGGCGTGTATGTGCCCTACACCCAG  
 R Q L S S T Y R D L R K G V Y V P Y T Q  
 GGCAAGTGGGAAGGGGAGCTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTC  
 G K W E G E L G T D L V S I P H G P N V  
 ACTGTGCGTGCCAAACATTGCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCC  
 T V R A N I A A I T E S D K F F I N G S  
 AACTGGGAAGGCATCCTGGGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTG  
 N W E G I L G L A Y A E I A R P D D S L  
 GAGCCTTTCTTTGACTCTCTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAG  
 E P F F D S L V K Q T H V P N L F S L Q  
 CTTTGTGGTGCTGGCTTCCCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTGCGAGGGAGC  
 L C G A G F P L N Q S E V L A S V G G S  
 ATGATCATTTGGAGGTATCGACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATC  
 M I I G G I D H S L Y T G S L W Y T P I  
 CGGCGGGAGTGGTATTATGAGGTATCATTGTGCGGGTGGAGATCAATGGACAGGATCTG  
 R R E W Y Y E V I I V R V E I N G Q D L  
 AAAATGGACTGCAAGGAGTACAACATATGACAAGAGCATTGTGGACAGTGGCACCACCAAC  
 K M D C K E Y N Y D K S I V D S G T T N  
 CTTGTTTGGCCAAGAAAGTGTGTTGAAGCTGCAGTCAAATCCATCAAGGCAGCCTCCTCC  
 L R L P K K V F E A A V K S I K A A S S  
 ACGGAGAAGTTCCCTGATGGTTTCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGC  
 T E K F P D G F W L G E Q L V C W Q A G  
 ACCACCCCTTGAACATTTTCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAAC  
 T T P W N I F P V I S L Y L M G E V T N

**FIGURE 7 (2)**

CAGTCCTTCCGCATCACCATCCTTCCGCAGCAATACCTGCGGCCAGTGGAAGATGTGGCC  
Q S F R I T I L P Q Q Y L R P V E D V A

ACGTCCCAAGACGACTGTTACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTTATG  
T S Q D D C Y K F A I S Q S S T G T V M

GGAGCTGTTATCATGGAGGGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAATTGGC  
G A V I M E G F Y V V F D R A R K R I G

TTTGCTGTCAGCGCTTGCCATGTGCACGATGAGTTCAGGACGGCAGCGGTGGAAGGCCCT  
F A V S A C H V H D E F R T A A V E G P

TTTGTCACCTTGACATGGAAGACTGTGGCTACAACATTCCACAGACAGATGAGTCATGA  
F V T L D M E D C G Y N I P Q T D E S \*

FIGURE 8 (1)

ATGACTCAGCATGGTATTTCGTCTGCCACTGCCGTAGCGGTCTGGGTGGTGTCTCCACTGGGT  
 M T Q H G I R L P L R S G L G G A P L G -  
 CTGCGTCTGCCCCGGGAGACCGACGAAGAGCCCCGAGGAGCCCCGGCCGGAGGGGCAGCTTT  
 L R L P R E T D E E P E E P G R R G S F -  
 GTGGAGATGGTGGACAACCTGAGGGGCAAGTCGGGGCAGGGCTACTACGTGGAGATGACC  
 V E M V D N L R G K S G Q G Y Y V E M T -  
 GTGGGCAGCCCCCGCAGACGCTCAACATCCTGGTGGATACAGGCAGCAGTAACCTTTGCA  
 V G S P P Q T L N I L V D T G S S N F A -  
 GTGGGTGCTGCCCCCACCCTTCCTGCATCGCTACTACCAGAGGCAGCTGTCCAGCACA  
 V G A A P H P F L H R Y Y Q R Q L S S T -  
 TACCGGGACCTCCGGAAGGGCGTGTATGTGCCCTACACCCAGGGCAAGTGGGAAGGGGAG  
 Y R D L R K G V Y V P Y T Q G K W E G E -  
 CTGGGCACCGACCTGGTAAGCATCCCCCATGGCCCCAACGTCACTGTGCGTGCCAAACATT  
 L G T D L V S I P H G P N V T V R A N I -  
 GCTGCCATCACTGAATCAGACAAGTTCTTCATCAACGGCTCCAACCTGGGAAGGCATCCTG  
 A A I T E S D K F F I N G S N W E G I L -  
 GGGCTGGCCTATGCTGAGATTGCCAGGCCTGACGACTCCCTGGAGCCTTTCTTTGACTCT  
 G L A Y A E I A R P D D S L E P F F D S  
 CTGGTAAAGCAGACCCACGTTCCCAACCTCTTCTCCCTGCAGCTTTGTGGTGTGGCTTC  
 L V K Q T H V P N L F S L Q L C G A G F -  
 CCCCTCAACCAGTCTGAAGTGCTGGCCTCTGTGCGAGGGAGCATGATCATTGGAGGTATC  
 P L N Q S E V L A S V G G S M I I G G I -  
 GACCACTCGCTGTACACAGGCAGTCTCTGGTATACACCCATCCGGCGGGAGTGGTATTAT  
 D H S L Y T G S L W Y T P I R R E W Y Y -  
 GAGGTCATCATTGTGCGGGTGGAGATCAATGGACAGGATCTGAAAATGGACTGCAAGGAG  
 E V I I V R V E I N G Q D L K M D C K E  
 TACAACATATGACAAGAGCATTGTGGACAGTGGCACCACCAACCTTCGTTTGCCCAAGAAA  
 Y N Y D K S I V D S G T T N L R L P K K -  
 GTGTTTGAAGCTGCAGTCAAATCCATCAAGGCAGCCTCCTCCACGGAGAAGTTCCTGAT  
 V F E A A V K S I K A A S S T E K F P D -  
 GGTTCCTGGCTAGGAGAGCAGCTGGTGTGCTGGCAAGCAGGCACCACCCCTTGGAACATT  
 G F W L G E Q L V C W Q A G T T P W N I -  
 TTCCAGTCATCTCACTCTACCTAATGGGTGAGGTTACCAACCAGTCCTTTTCGCATCACC  
 F P V I S L Y L M G E V T N Q S F R I T -  
 ATCCTTCCGCAGCAATACCTGCGGCCAGTGGGAAGATGTGGCCACGTCCCAAGACGACTGT  
 I L P Q Q Y L R P V E D V A T S Q D D C -

**FIGURE 8 (2)**

TACAAGTTTGCCATCTCACAGTCATCCACGGGCACTGTTATGGGAGCTGTTATCATGGAG  
Y K F A I S Q S S T G T V M G A V I M E -

GGCTTCTACGTTGTCTTTGATCGGGCCCGAAAACGAATTGGCTTTGCTGTCAGCGCTTGC  
G F Y V V F D R A R K R I G F A V S A C -

CATTAG

H \*

**FIGURE 9**

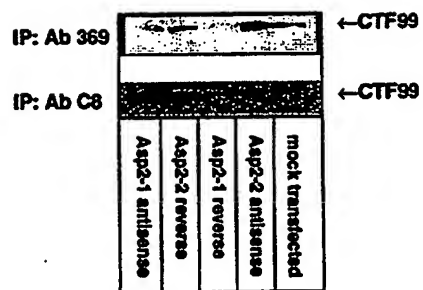
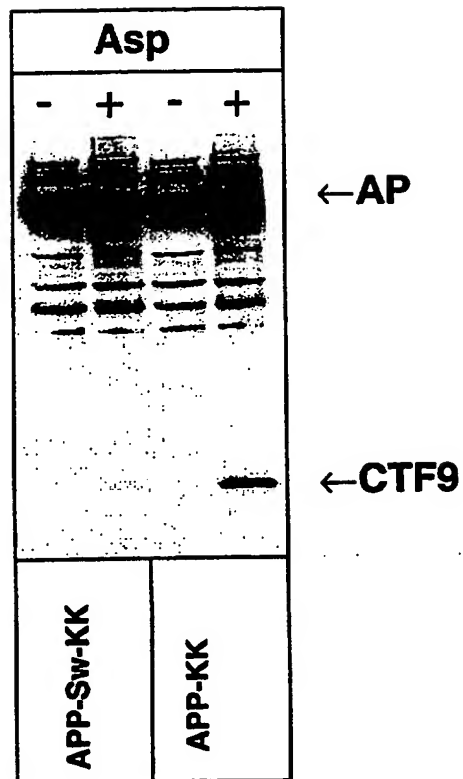


FIGURE 10



**FIGURE 11**

MAOALPWLLLLWMGAGVLPAHGTQHGI RLPLRSGLGGA PLGLRLPRETDEE  
PEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPQTLN ILVDTGSSNFA  
VGAAPHPFLHRYYQRQLSSTYRDLRKGVVVPYTQGWEGELGTDLVSI PH  
GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS  
LVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLW  
YTPIRREWYVEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLR LPKK  
VFEAAVKSIIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG  
EVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIME  
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQT  
DES

**FIGURE 12**

MAOALPWLLLWMGAGVLPAHGTQHGI~~RLPL~~RSGLGGAPLGLRLPRETDEE  
PEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFA  
VGAAPHPFLHRYYQRQLSSTYRDLRKG~~VYVP~~YTQGWEGELGTDLVSI~~PH~~  
GPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDS  
LVKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLW  
YTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRLPKK  
VFEEAVKSIKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMG  
EVTNQSF~~RITILPQOYL~~RPVEDVATSQDDCYKFAISQSSTGTVMGAVIME  
GFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQT  
DESHHHHHH



**Alzheimer's Disease Secretase, APP Substrates Therefor, and Uses Therefor**

The present application is a continuation-in-part of United States Patent Application 09/416,901, filed October 13, 1999 which claims priority benefit of United States Provisional Patent Application No. 60/155,493, filed September 23, 1999 and United States Provisional Patent Application 60/169232, filed December 6, 1999. The present application also claims priority benefit as a continuation-in-part of United States Patent Application Serial No. 09/404,133 and PCT/US99/20881, both filed September 23, 1999, both of which in turn claim priority benefit of United States Provisional Patent Application No. 60/101,594, filed September 24, 1998. All of these priority applications are hereby incorporated by reference in their entirety.

**FIELD OF THE INVENTION**

The present invention relates to Alzheimer's Disease, amyloid protein precursor, amyloid beta peptide, and human aspartyl proteases, as well as a method for the identification of agents that modulate the activity of these polypeptides and thereby are candidates to modulate the progression of Alzheimer's disease.

**BACKGROUND OF THE INVENTION**

Alzheimer's disease (AD) causes progressive dementia with consequent formation of amyloid plaques, neurofibrillary tangles, gliosis and neuronal loss. The disease occurs in both genetic and sporadic forms whose clinical course and pathological features are quite similar. Three genes have been discovered to date which, when mutated, cause an autosomal dominant form of Alzheimer's disease. These encode the amyloid protein precursor (APP) and two related proteins, presenilin-1 (PS1) and presenilin-2 (PS2), which, as their names suggest, are structurally and functionally related. Mutations in any of the three proteins have been observed to enhance proteolytic processing of APP via an intracellular pathway that produces amyloid beta peptide (A $\beta$  peptide, or sometimes here as Abeta), a 40-42 amino acid long peptide that is the primary component of amyloid plaque in AD.

Dysregulation of intracellular pathways for proteolytic processing may be central to the pathophysiology of AD. In the case of plaque formation, mutations in APP, PS1 or PS2 consistently alter the proteolytic processing of APP so as to enhance

formation of A $\beta$  1-42, a form of the A $\beta$  peptide which seems to be particularly amyloidogenic, and thus very important in AD. Different forms of APP range in size from 695-770 amino acids, localize to the cell surface, and have a single C-terminal transmembrane domain. Examples of specific isotypes of APP which are currently known to exist in humans are the 695-amino acid polypeptide described by Kang *et al.* (1987), *Nature* 325: 733-736 which is designated as the "normal" APP; the 751 amino acid polypeptide described by Ponte *et al.* (1988), *Nature* 331: 525-527 (1988) and Tanzi *et al.* (1988), *Nature* 331: 528-530; and the 770 amino acid polypeptide described by Kitaguchi *et al.* (1988), *Nature* 331: 530-532. The Abeta peptide is derived from a region of APP adjacent to and containing a portion of the transmembrane domain. Normally, processing of APP at the  $\alpha$ -secretase site cleaves the midregion of the A $\beta$  sequence adjacent to the membrane and releases the soluble, extracellular domain of APP from the cell surface. This  $\alpha$ -secretase APP processing creates soluble APP-  $\alpha$ , (sAPP $\alpha$ ) which is normal and not thought to contribute to AD.

Pathological processing of APP at the  $\beta$ - and  $\gamma$ -secretase sites, which are located N-terminal and C-terminal to the  $\alpha$ -secretase site, respectively, produces a very different result than processing at the  $\alpha$  site. Sequential processing at the  $\beta$ - and  $\gamma$ -secretase sites releases the A $\beta$  peptide, a peptide possibly very important in AD pathogenesis. Processing at the  $\beta$ - and  $\gamma$ -secretase sites can occur in both the endoplasmic reticulum (in neurons) and in the endosomal/lysosomal pathway after reinternalization of cell surface APP (in all cells). Despite intense efforts, for 10 years or more, to identify the enzymes responsible for processing APP at the  $\beta$  and  $\gamma$  sites, to produce the A $\beta$  peptide, those proteases remained unknown until this disclosure.

## SUMMARY OF THE INVENTION

Here, for the first time, we report the identification and characterization of the  $\beta$  secretase enzyme, termed Aspartyl Protease 2 (Asp2). We disclose some known and some novel human aspartic proteases that can act as  $\beta$ -secretase proteases and, for

the first time, we explain the role these proteases have in AD. We describe regions in the proteases critical for their unique function and for the first time characterize their substrate. This is the first description of expressed isolated purified active protein of this type, assays that use the protein, in addition to the identification and creation of useful cell lines and inhibitors. We also identify and characterize both  $\alpha$ -secretase and  $\beta$ -secretase activities of a protease, designated as Asp1.

Here we disclose a number of variants of the Asp2 gene and peptide.

In one aspect, the invention provides any isolated or purified nucleic acid polynucleotide that codes for a protease capable of cleaving the beta ( $\beta$ ) secretase cleavage site of APP that contains two or more sets of special nucleic acids, where the special nucleic acids are separated by nucleic acids that code for about 100 to 300 amino acid positions, where the amino acids in those positions may be any amino acids, where the first set of special nucleic acids consists of the nucleic acids that code for the peptide DTG, where the first nucleic acid of the first special set of nucleic acids is the first special nucleic acid, and where the second set of nucleic acids code for either the peptide DSG or DTG, where the last nucleic acid of the second set of nucleic acids is the last special nucleic acid, with the proviso that the nucleic acids disclosed in SEQ ID NO. 1 and SEQ ID NO. 3 are not included. In a preferred embodiment, the two sets of special nucleic acids are separated by nucleic acids that code for about 125 to 222 amino acid positions, which may be any amino acids. In a highly preferred embodiment, the two sets of special nucleic acids are separated by nucleic acids that code for about 150 to 196, or 150-190, or 150 to 172 amino acid positions, which may be any amino acids. In a particular preferred embodiment, the two sets are separated by nucleic acids that code for about 172 amino acid positions, which may be any amino acids. An exemplary nucleic acid polynucleotide comprises the acid nucleotide sequence in SEQ ID NO. 5. In another particular preferred embodiment, the two sets are separated by nucleic acids that code for about 196 amino acids. An exemplary polynucleotide comprises the nucleotide sequence in SEQ ID NO. 5. In another particular embodiment, the two sets of nucleotides are separated by nucleic acids that code for about 190 amino acids. An exemplary polynucleotide

comprises the nucleotide sequence in SEQ ID NO. 1. Preferably, the first nucleic acid of the first special set of amino acids, that is, the first special nucleic acid, is operably linked to any codon where the nucleic acids of that codon codes for any peptide comprising from 1 to 10,000 amino acid (positions). In one variation, the first special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: any reporter proteins or proteins which facilitate purification. For example, the first special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: immunoglobulin-heavy chain, maltose binding protein, glutathione S transferase, Green Fluorescent protein, and ubiquitin. In another variation, the last nucleic acid of the second set of special amino acids, that is, the last special nucleic acid, is operably linked to nucleic acid polymers that code for any peptide comprising any amino acids from 1 to 10,000 amino acids. In still another variation, the last special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: any reporter proteins or proteins which facilitate purification. For example, the last special nucleic acid is operably linked to nucleic acid polymers that code for any peptide selected from the group consisting of: immunoglobulin-heavy chain, maltose binding protein, glutathione S transferase, Green Fluorescent protein, and ubiquitin.

In a related aspect, the invention provides any isolated or purified nucleic acid polynucleotide that codes for a protease capable of cleaving the beta secretase cleavage site of APP that contains two or more sets of special nucleic acids, where the special nucleic acids are separated by nucleic acids that code for about 100 to 300 amino acid positions, where the amino acids in those positions may be any amino acids, where the first set of special nucleic acids consists of the nucleic acids that code for DTG, where the first nucleic acid of the first special set of nucleic acids is the first special nucleic acid, and where the second set of nucleic acids code for either DSG or DTG, where the last nucleic acid of the second set of special nucleic acids is the last special nucleic acid, where the first special nucleic acid is operably linked to nucleic acids that code for any number of amino acids from zero to 81 amino acids and where

each of those codons may code for any amino acid. In a preferred embodiment, the first special nucleic acid is operably linked to nucleic acids that code for any number of from 64 to 77 amino acids where each codon may code for any amino acid. In a particular embodiment, the first special nucleic acid is operably linked to nucleic acids that code for 71 amino acids. For example, the first special nucleic acid is operably linked to 71 amino acids and where the first of those 71 amino acids is the amino acid T. In a preferred embodiment, the polynucleotide comprises a sequence that is at least 95% identical to a human Asp1 or Asp2 sequence as taught herein. In another preferred embodiment, the first special nucleic acid is operably linked to nucleic acids that code for any number of from 30 to 54 amino acids, or 35 to 47 amino acids, or 40 to 54 amino acids where each codon may code for any amino acid. In a particular embodiment, the first special nucleic acid is operably linked to nucleic acids that code for 47 amino acids. For example, the first special nucleic acid is operably linked to 47 codons where the first of those 47 amino acids is the amino acid E.

In another related aspect, the invention provides for any isolated or purified nucleic acid polynucleotide that codes for a protease capable of cleaving the beta ( $\beta$ ) secretase cleavage site of APP and that contains two or more sets of special nucleic acids, where the special nucleic acids are separated by nucleic acids that code for about 100 to 300 amino acid positions, where the amino acids in those positions may be any amino acids, where the first set of special nucleic acids consists of the nucleic acids that code for the peptide DTG, where the first nucleic acid of the first special set of amino acids is, the first special nucleic acid, and where the second set of special nucleic acids code for either the peptide DSG or DTG, where the last nucleic acid of the second set of special nucleic acids, the last special nucleic acid, is operably linked to nucleic acids that code for any number of codons from 50 to 170 codons. In a preferred embodiment, the last special nucleic acid is operably linked to nucleic acids comprising from 100 to 170 codons. In a highly preferred embodiment, the last special nucleic acid is operably linked to nucleic acids comprising from 142 to 163 codons. In a particular embodiment, the last special nucleic acid is operably linked to nucleic acids comprising about 142 codons, or about 163 codons, or about 170

codons. In a highly preferred embodiment, the polynucleotide comprises a sequence that is at least 95% identical to aspartyl-protease encoding sequences taught herein. In one variation, the second set of special nucleic acids code for the peptide DSG. In another variation, the first set of nucleic acid polynucleotide is operably linked to a peptide purification tag. For example, the nucleic acid polynucleotide is operably linked to a peptide purification tag which is six histidine. In still another variation, the first set of special nucleic acids are on one polynucleotide and the second set of special nucleic acids are on a second polynucleotide, where both first and second polynucleotides have at least 50 codons. In one embodiment of this type, both of the polynucleotides are in the same solution. In a related aspect, the invention provides a vector which contains a polynucleotide as described above, or a cell or cell line which is transformed or transfected with a polynucleotide as described above or with a vector containing such a polynucleotide.

In still another aspect, the invention provides an isolated or purified peptide or protein comprising an amino acid polymer that is a protease capable of cleaving the beta ( $\beta$ ) secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are separated by about 100 to 300 amino acid positions, where each amino acid position can be any amino acid, where the first set of special amino acids consists of the peptide DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, where the second set of amino acids is selected from the peptide comprising either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, with the proviso that the proteases disclosed in SEQ ID NO. 2 and SEQ ID NO. 4 are not included. In preferred embodiments, the two sets of amino acids are separated by about 125 to 222 amino acid positions or about 150 to 196 amino acids, or about 150-190 amino acids, or about 150 to 172 amino acids, where in each position it may be any amino acid. In a particular embodiment, the two sets of amino acids are separated by about 172 amino acids. For example, the protease has the amino acid sequence described in SEQ ID NO 6. In another particular embodiment, the two sets of amino acids are separated by about 196 amino acids. For example, the two sets of amino

acids are separated by the same amino acid sequences that separate the same set of special amino acids in SEQ ID NO 4. In another particular embodiment, the two sets of nucleotides are separated by about 190 amino acids. For example, the two sets of nucleotides are separated by the same amino acid sequences that separate the same set of special amino acids in SEQ ID NO 2. In one embodiment, the first amino acid of the first special set of amino acids, that is, the first special amino acid, is operably linked to any peptide comprising from 1 to 10,000 amino acids. In another embodiment, the first special amino acid is operably linked to any peptide selected from the group consisting of: any reporter proteins or proteins which facilitate purification. In particular embodiments, the first special amino acid is operably linked to any peptide selected from the group consisting of: immunoglobulin-heavy chain, maltose binding protein, glutathione S transferase, Green Fluorescent protein, and ubiquitin. In still another variation, the last amino acid of the second set of special amino acids, that is, the last special amino acid, is operably linked to any peptide comprising any amino acids from 1 to 10,000 amino acids. By way of nonlimiting example, the last special amino acid is operably linked any peptide selected from the group consisting of any reporter proteins or proteins which facilitate purification. In particular embodiments, the last special amino acid is operably linked to any peptide selected from the group consisting of: immunoglobulin-heavy chain, maltose binding protein, glutathione S transferase, Green Fluorescent protein, and ubiquitin.

In a related aspect, the invention provides any isolated or purified peptide or protein comprising an amino acid polypeptide that codes for a protease capable of cleaving the beta secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are separated by about 100 to 300 amino acid positions, where each amino acid in each position can be any amino acid, where the first set of special amino acids consists of the amino acids DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, D, and where the second set of amino acids is either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, G, where the first special amino acid is operably linked to amino acids that code for any

number of amino acids from zero to 81 amino acid positions where in each position it may be any amino acid. In a preferred embodiment, the first special amino acid is operably linked to a peptide from about 30-77 or about 64 to 77 amino acids positions where each amino acid position may be any amino acid. In a particular embodiment, the first special amino acid is operably linked to a peptide 35, 47, 71, or 77 amino acids. In a very particular embodiment, the first special amino acid is operably linked to 71 amino acids and the first of those 71 amino acids is the amino acid T. For example, the polypeptide comprises a sequence that is at least 95% identical to an aspartyl protease sequence as described herein. In another embodiment, the first special amino acid is operably linked to any number of from 40 to 54 amino acids (positions) where each amino acid position may be any amino acid. In a particular embodiment, the first special amino acid is operably linked to amino acids that code for a peptide of 47 amino acids. In a very particular embodiment, the first special amino acid is operably linked to a 47 amino acid peptide where the first those 47 amino acids is the amino acid E. In another particular embodiment, the first special amino acid is operably linked to the same corresponding peptides from SEQ ID NO. 3 that are 35, 47, 71, or 77 peptides in length, beginning counting with the amino acids on the first special sequence, DTG, towards the N-terminal of SEQ ID NO. 3. In another particular embodiment, the polypeptide comprises a sequence that is at least 95% identical to the same corresponding amino acids in SEQ ID NO. 4, that is, identical to that portion of the sequences in SEQ ID NO. 4, including all the sequences from both the first and or the second special nucleic acids, toward the N terminal, through and including 71, 47, 35 amino acids before the first special amino acids. For example, the complete polypeptide comprises the peptide of 71 amino acids, where the first of the amino acid is T and the second is Q.

In still another related aspect, the invention provides any isolated or purified amino acid polypeptide that is a protease capable of cleaving the beta ( $\beta$ ) secretase cleavage site of APP that contains two or more sets of special amino acids, where the special amino acids are separated by about 100 to 300 amino acid positions, where each amino acid in each position can be any amino acid, where the first set of special



amino acids consists of the amino acids that code for DTG, where the first amino acid of the first special set of amino acids is, the first special amino acid, D, and where the second set of amino acids are either DSG or DTG, where the last amino acid of the second set of special amino acids is the last special amino acid, G, which is operably linked to any number of amino acids from 50 to 170 amino acids, which may be any amino acids. In preferred embodiments, the last special amino acid is operably linked to a peptide of about 100 to 170 amino acids or about 142-163 amino acids. In particular embodiments, the last special amino acid is operably linked to a peptide of about 142 amino acids, or about 163 amino acids, or about 170 amino acids. For example, the polypeptide comprises a sequence that is at least 95% identical (and preferably 100% identical) to an aspartyl protease sequence as described herein. In one particular embodiment, the second set of special amino acids is comprised of the peptide with the amino acid sequence DSG. Optionally, the amino acid polypeptide is operably linked to a peptide purification tag, such as purification tag which is six histidine. In one variation, the first set of special amino acids are on one polypeptide and the second set of special amino acids are on a second polypeptide, where both first and second polypeptide have at least 50 amino acids, which may be any amino acids. In one embodiment of this type, both of the polypeptides are in the same vessel. The invention further includes a process of making any of the polynucleotides, vectors, or cells described herein; and a process of making any of the polypeptides described herein.

In yet another related aspect, the invention provides a purified polynucleotide comprising a nucleotide sequence that encodes a polypeptide having aspartyl protease activity, wherein the polypeptide has an amino acid sequence characterized by: (a) a first tripeptide sequence DTG; (b) a second tripeptide sequence selected from the group consisting of DSG and DTG; and (c) about 100 to 300 amino acids separating the first and second tripeptide sequences, wherein the polypeptide cleaves the beta secretase cleavage site of amyloid protein precursor. In one embodiment, the polypeptide comprises an amino acid sequence depicted in SEQ ID NO: 2 or 4, whereas in another embodiment, the polypeptide comprises an amino acid sequence

other than the amino acid sequences set forth in SEQ ID NOs: 2 and 4. Similarly, the invention provides a purified polynucleotide comprising a nucleotide sequence that encodes a polypeptide that cleaves the beta secretase cleavage site of amyloid protein precursor; wherein the polynucleotide includes a strand that hybridizes to one or more of SEQ ID NOs: 3, 5, and 7 under the following hybridization conditions:

hybridization overnight at 42°C for 2.5 hours in 6 X SSC/0.1% SDS, followed by washing in 1.0 X SSC at 65°C, 0.1% SDS. In one embodiment, the polypeptide comprises an amino acid sequence depicted in SEQ ID NO: 2 or 4, whereas in another embodiment, the polypeptide comprises an amino acid sequence other than the amino acid sequences set forth in SEQ ID NOs: 2 and 4. Likewise, the invention provides a purified polypeptide having aspartyl protease activity, wherein the polypeptide is encoded by polynucleotides as described in the preceding sentences. The invention also provides a vector or host cell comprising such polynucleotides, and a method of making the polypeptides using the vectors or host cells to recombinantly express the polypeptide.

In yet another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide, said polynucleotide encoding a Hu-Asp polypeptide and having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a nucleotide sequence encoding a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), wherein said Hu-Asp1, Hu-Asp2(a) and Hu-Asp2(b) polypeptides have the complete amino acid sequence of SEQ ID NO. 2, SEQ ID NO. 4, and SEQ ID NO. 6, respectively; and

(b) a nucleotide sequence complementary to the nucleotide sequence of (a).

Several species are particularly contemplated. For example, the invention provides a nucleic acid and molecule wherein said Hu-Asp polypeptide is Hu-Asp1, and said polynucleotide molecule of 1(a) comprises the nucleotide sequence of SEQ ID NO. 1; and a nucleic acid molecule wherein said Hu-Asp polypeptide is Hu-Asp2(a), and said polynucleotide molecule of 1(a) comprises the nucleotide

sequence of SEQ ID NO. 4; and a nucleic acid molecule wherein said Hu-Asp polypeptide is Hu-Asp2(b), and said polynucleotide molecule of 1(a) comprises the nucleotide sequence of SEQ ID NO. 5. In addition to the foregoing, the invention provides an isolated nucleic acid molecule comprising polynucleotide which hybridizes under stringent conditions to a polynucleotide having the nucleotide sequence in (a) or (b) as described above.

Additionally, the invention provides a vector comprising a nucleic acid molecule as described in the preceding paragraph. In a preferred embodiment, the nucleic acid molecule is operably linked to a promoter for the expression of a Hu-Asp polypeptide. Individual vectors which encode Hu-Asp1, and Hu-Asp2(a), and Hu-Asp2(b) are all contemplated. Likewise, the invention contemplates a host cell comprising any of the foregoing vectors, as well as a method of obtaining a Hu-Asp polypeptide comprising culturing such a host cell and isolating the Hu-Asp polypeptide. Host cells of the invention include bacterial cells, such as *E. coli*, and eukaryotic cells. Among the eukaryotic cells that are contemplated are insect cells, such as sf9 or High 5 cells; and mammalian cells, such as human, rodent, lagomorph, and primate. Preferred human cells include HEK293, and IMR-32 cells. Other preferred mammalian cells include COS-7, CHO-K1, Neuro-2A, and 3T3 cells. Also among the eukaryotic cells that are contemplated are a yeast cell and an avian cell.

In a related aspect, the invention provides an isolated Hu-Asp1 polypeptide comprising an amino acid sequence at least 95% identical to a sequence comprising the amino acid sequence of SEQ ID NO. 2. The invention also provides an isolated Hu-Asp2(a) polypeptide comprising an amino acid sequence at least 95% identical to a sequence comprising the amino acid sequence of SEQ ID NO. 4. The invention also provides an isolated Hu-Asp2(a) polypeptide comprising an amino acid sequence at least 95% identical to a sequence comprising the amino acid sequence of SEQ ID NO. 8.

In still another aspect, the invention provides an isolated antibody that binds specifically to any Hu-Asp polypeptide described herein, especially the polypeptide described in the preceding paragraphs.

The invention also provides several assays involving aspartyl protease enzymes of the invention. For example, the invention provides

a method to identify a cell that can be used to screen for inhibitors of  $\beta$  secretase activity comprising:

(a) identifying a cell that expresses a protease capable of cleaving APP at the  $\beta$  secretase site, comprising:

- i) collect the cells or the supernatant from the cells to be identified
- ii) measure the production of a critical peptide, where the critical peptide is selected from the group consisting of either the APP C-terminal peptide or soluble APP,
- iii) select the cells which produce the critical peptide.

In one variation, the cells are collected and the critical peptide is the APP C-terminal peptide created as a result of the  $\beta$  secretase cleavage. In another variation, the supernatant is collected and the critical peptide is soluble APP, where the soluble APP has a C-terminus created by  $\beta$  secretase cleavage. In preferred embodiments, the cells contain any of the nucleic acids or polypeptides described above and the cells are shown to cleave the  $\beta$ -secretase site of any peptide having the following peptide structure, P2, P1, P1', P2', where P2 is K or N, where P1 is M or L, where P1' is D, where P2' is A. The method where P2 is K and P1 is M. The method where P2 is N and P1 is L.

In still another aspect, the invention provides novel isoforms of amyloid protein precursor (APP) where the last two carboxy terminus amino acids of that isoform are both lysine residues. In this context, the term "isoform" is defined as any APP polypeptide, including APP variants (including mutations), and APP fragments that exists in humans, such as those described in US 5,766,846, col 7, lines 45-67, incorporated into this document by reference, modified as described herein by the inclusion of two C-terminal lysine residues. For example, the invention provides a polypeptide comprising the isoform known as APP695, modified to include two lysine residues as its last two carboxy terminus amino acids. An exemplary polypeptide

comprises the amino acid sequence set forth in SEQ ID NO. 16. The invention further includes APP isoform variants as set forth in SEQ ID NOs. 18 and 20. The invention further includes all polynucleotides that encode an APP protein that has been modified to include two C-terminal lysines; as well as any eukaryotic cell line comprising such nucleic acids or polypeptides. Preferred cell lines include a mammalian cell line (*e.g.*, HEK293, Neuro2a).

Thus, in one embodiment, the invention provides a polypeptide comprising the amino acid sequence of a mammalian amyloid protein precursor (APP) or fragment thereof containing an APP cleavage site recognizable by a mammalian  $\beta$ -secretase, and further comprising two lysine residues at the carboxyl terminus of the amino acid sequence of the mammalian APP or APP fragment. As taught herein in detail, the addition of two additional lysine residues to APP sequences has been found to greatly increase  $A\beta$  processing of the APP in APP processing assays. Thus, the di-lysine modified APP reagents of the invention are particularly useful in assays to identify modulators of  $A\beta$  production, for use in designing therapeutics for the treatment or prevention of Alzheimer's disease. In one embodiment, the polypeptide comprises the complete amino acid sequence of a mammalian amyloid protein precursor (APP), and further comprises the two lysine residues at the carboxyl terminus of the amino acid sequence of the mammalian amyloid protein precursor. In an alternative embodiment, the polypeptide comprises only a fragment of the APP, the fragment containing at least that portion of APP that is cleaved by a mammalian  $\beta$ -secretase (or  $\alpha$ -secretase or  $\gamma$ -secretase) in the formation of  $A\beta$  peptides.

The practice of assays that monitor cleavage of APP can be facilitated by attaching a marker to a portion of the APP. Measurement of retained or liberated marker can be used to quantitate the amount of APP cleavage that occurs in the assay, *e.g.*, in the presence or absence of a putative modulator of cleavage activity. Thus, in one preferred embodiment, the polypeptide of the invention further includes a marker. For example, the marker comprises a reporter protein amino acid sequence attached to the APP amino acid sequence. Exemplary reporter proteins include a fluorescing protein (*e.g.*, green fluorescing proteins, luciferase) or an enzyme that is used to cleave

a substrate to produce a colorimetric cleavage product. Also contemplated are tag sequences which are commonly used as epitopes for quantitative immunoassays.

In a preferred embodiment, the di-lysine-modified APP of the invention is a human APP. For example, human APP isoforms such as APP695, APP751, and APP770, modified to include the two lysines, are contemplated. In a preferred embodiment, the APP isoform comprises at least one variation selected from the group consisting of a Swedish KM→NL mutation and a London V717→F mutation, or any other mutation that has been observed in a subpopulation that is particularly prone to development of Alzheimer's disease. These mutations are recognized as mutations that influence APP processing into A $\beta$ . In a highly preferred embodiment, the APP protein or fragment thereof comprises the APP-Sw  $\beta$ -secretase peptide sequence NLDA, which is associated with increased levels of A $\beta$  processing and therefore is particularly useful in assays relating to Alzheimer's research. More particularly, the APP protein or fragment thereof preferably comprises the APP-Sw  $\beta$ -secretase peptide sequence SEVNLD AEFR (SEQ ID NO: 63).

In one preferred embodiment, the APP protein or fragment thereof further includes an APP transmembrane domain carboxy-terminal to the APP-Sw  $\beta$ -secretase peptide sequence. Polypeptides that include the TM domain are particularly useful in cell-based APP processing assays. In contrast, embodiments lacking the TM domain are useful in cell-free assays of APP processing.

In addition to working with APP from humans and various animal models, researchers in the field of Alzheimer's also have construct chimeric APP polypeptides which include stretches of amino acids from APP of one species (e.g., humans) fused to stretches of APP from one or more other species (e.g., rodent). Thus, in another embodiment of the polypeptide of the invention, the APP protein or fragment thereof comprises a chimeric APP, the chimeric APP including partial APP amino acid sequences from at least two species. A chimeric APP that includes amino acid sequence of a human APP and a rodent APP is particularly contemplated.

In a related aspect, the invention provides a polynucleotide comprising a nucleotide sequence that encodes a polypeptide as described in the preceding

paragraphs. Such a polynucleotide is useful for recombinant expression of the polypeptide of the invention for use in APP processing assays. In addition, the polynucleotide is useful for transforming into cells to produce recombinant cells that express the polypeptide of the invention, which cells are useful in cell-based assays to identify modulators of APP processing. Thus, in addition to polynucleotides, the invention provides a vector comprising such polynucleotides, especially expression vectors where the polynucleotide is operably linked to a promoter to promote expression of the polypeptide encoded by the polynucleotide in a host cell. The invention further provides a host cell transformed or transfected with such a polynucleotide or a vector. Among the preferred host cells are mammalian cells, especially human cells.

In another, related embodiment, the invention provides a polypeptide useful for assaying for modulators of  $\beta$ -secretase activity, said polypeptide comprising an amino acid sequence of the formula  $\text{NH}_2\text{-X-Y-Z-KK-COOH}$ ; wherein X, Y, and Z each comprise an amino acid sequence of at least one amino acid; wherein  $\text{NH}_2\text{-X}$  comprises an amino-terminal amino acid sequence having at least one amino acid residue; wherein Y comprises an amino acid sequence of a  $\beta$ -secretase recognition site of a mammalian amyloid protein precursor (APP); and wherein  $\text{Z-KK-COOH}$  comprises a carboxy-terminal amino acid sequence ending in two lysine (K) residues.

In one preferred variation, the carboxyl-terminal amino acid sequence Z includes a hydrophobic domain that is a transmembrane domain in host cells that express the polypeptide. Host cells that express such a polypeptide are particularly useful in assays described herein for identifying modulators of APP processing. In another preferred variation, the amino-terminal amino acid sequence X includes an amino acid sequence of a reporter or marker protein, as described above. In still another preferred variation, the  $\beta$ -secretase recognition site Y comprises the human APP-Sw  $\beta$ -secretase peptide sequence NLDA. It will be apparent that these preferred variations are not mutually exclusive of each other -- they may be combined in a single polypeptide. The invention further provides a polynucleotide comprising a nucleotide sequence that

encodes such polypeptides, vectors which comprise such polynucleotides, and host cells which comprises such vectors, polynucleotides, and/or polypeptides.

In yet another aspect, the invention provides a method for identifying inhibitors of an enzyme that cleaves the beta secretase cleavable site of APP comprising:

- a) culturing cells in a culture medium under conditions in which the enzyme causes processing of APP and release of amyloid beta-peptide into the medium and causes the accumulation of CTF99 fragments of APP in cell lysates,
- b) exposing the cultured cells to a test compound; and specifically determining whether the test compound inhibits the function of the enzyme by measuring the amount of amyloid beta-peptide released into the medium and/or the amount of CTF99 fragments of APP in cell lysates;
- c) identifying test compounds diminishing the amount of soluble amyloid beta peptide present in the culture medium and diminution of CTF99 fragments of APP in cell lysates as Asp2 inhibitors. In preferred embodiments, the cultured cells are a human, rodent or insect cell line. It is also preferred that the human or rodent cell line exhibits  $\beta$  secretase activity in which processing of APP occurs with release of amyloid beta-peptide into the culture medium and accumulation of CTF99 in cell lysates. Among the contemplated test compounds are antisense oligomers directed against the enzyme that exhibits  $\beta$  secretase activity, which oligomers reduce release of soluble amyloid beta-peptide into the culture medium and accumulation of CTF99 in cell lysates.

In yet another aspect, the invention provides a method for the identification of an agent that decreases the activity of a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), the method comprising:

- a) determining the activity of said Hu-Asp polypeptide in the presence of a test agent and in the absence of a test agent; and
- b) comparing the activity of said Hu-Asp polypeptide determined in the presence of said test agent to the activity of said Hu-Asp polypeptide determined in the absence of said test agent; whereby a lower level of activity in the presence of said



test agent than in the absence of said test agent indicates that said test agent has decreased the activity of said Hu-Asp polypeptide.

In a related aspect, the invention provides a method for assaying for modulators of  $\beta$ -secretase activity, comprising the steps of:

(a) contacting a first composition with a second composition both in the presence and in the absence of a putative modulator compound, wherein the first composition comprises a mammalian  $\beta$ -secretase polypeptide or biologically active fragment thereof, and wherein the second composition comprises a substrate polypeptide having an amino acid sequence comprising a  $\beta$ -secretase cleavage site; (b) measuring cleavage of the substrate polypeptide in the presence and in the absence of the putative modulator compound; and (c) identifying modulators of  $\beta$ -secretase activity from a difference in cleavage in the presence versus in the absence of the putative modulator compound. A modulator that is a  $\beta$ -secretase antagonist (inhibitor) reduces such cleavage, whereas a modulator that is a  $\beta$ -secretase agonist increases such cleavage. Since such assays are relevant to development of Alzheimer's disease therapeutics for humans, it will be readily apparent that, in one preferred embodiment, the first composition comprises a purified human Asp2 polypeptide. In one variation, the first composition comprises a soluble fragment of a human Asp2 polypeptide that retains Asp2  $\beta$ -secretase activity. Several such fragments (including  $\Delta$ TM fragments) are described herein in detail. Thus, in a particular embodiment, the soluble fragment is a fragment lacking an Asp2 transmembrane domain. Assaying to identify inhibitors of Asp1  $\beta$ -secretase activity also is contemplated.

The  $\beta$ -secretase cleavage site in APP is known, and it will be appreciated that the assays of the invention can be performed with either intact APP or fragments or analogs of APP that retain the  $\beta$ -secretase recognition and cleavage site. Thus, in one variation, the substrate polypeptide of the second composition comprises the amino acid sequence SEVNLDAEFR, which includes the  $\beta$ -secretase recognition site of human APP that contains the "Swiss" mutation. In another variation, the substrate polypeptide of the second composition comprises the amino acid sequence EVKMDAEF. In another variation, the second composition comprises a polypeptide

having an amino acid sequence of a human amyloid precursor protein (APP). For example, the human amyloid precursor protein is selected from the group consisting of: APP695, APP751, and APP770. Preferably, the human amyloid precursor protein (irrespective of isoform selected) includes at least one mutation selected from a KM→NL Swiss mutation and a V→F London mutation. As explained elsewhere, one preferred embodiment involves a variation wherein the polypeptide having an amino acid sequence of a human APP further comprises an amino acid sequence comprising a marker sequence attached amino-terminal to the amino acid sequence of the human amyloid precursor protein. Preferably, the polypeptide having an amino acid sequence of a human APP further comprises two lysine residues attached to the carboxyl terminus of the amino acid sequence of the human APP. The assays can be performed in a cell free setting, using cell-free enzyme and cell-free substrate, or can be performed in a cell-based assay wherein the second composition comprises a eukaryotic cell that expresses amyloid precursor protein (APP) or a fragment thereof containing a  $\beta$ -secretase cleavage site. Preferably, the APP expressed by the host cell is an APP variant that includes two carboxyl-terminal lysine residues. It will also be appreciated that the  $\beta$ -secretase enzyme can be an enzyme that is expressed on the surface of the same cells.

The present invention provides isolated nucleic acid molecules comprising a polynucleotide that codes for a polypeptide selected from the group consisting of human aspartyl proteases. In particular, human aspartyl protease 1 (Hu-Asp1) and two alternative splice variants of human aspartyl protease-2 (Hu-Asp2), a "long" (L) form designated herein as Hu-Asp2(a) and a "short" (S) form designated Hu-Asp2(b). As used herein, all references to "Hu-Asp" should be understood to refer to all of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b). In addition, as used herein, all references to "Hu-Asp2" should be understood to refer to both Hu-Asp2(a) and Hu-Asp2(b). Hu-Asp1 is expressed most abundantly in pancreas and prostate tissues, while Hu-Asp2(a) and Hu-Asp2(b) are expressed most abundantly in pancreas and brain tissues. The invention also provides isolated Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides, as well as fragments thereof which exhibit aspartyl protease activity.

In a preferred embodiment, the nucleic acid molecules comprise a polynucleotide having a nucleotide sequence selected from the group consisting of residues 1-1554 of SEQ ID NO. 1, encoding Hu-Asp1, residues 1-1503 of SEQ ID NO. 3, encoding Hu-Asp2(a), and residues 1-1428 of SEQ ID NO.5, encoding Hu-Asp2(b). In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent conditions to a polynucleotide encoding Hu-Asp1, Hu-Asp2(a), Hu-Asp-2(b), or fragments thereof. European patent application EP 0 848 062 discloses a polypeptide referred to as "Asp 1," that bears substantial homology to Hu-Asp1, while international application WO 98/22597 discloses a polypeptide referred to as "Asp 2," that bears substantial homology to Hu-Asp2(a).

The present invention also provides vectors comprising the isolated nucleic acid molecules of the invention, host cells into which such vectors have been introduced, and recombinant methods of obtaining a Hu-Asp1, Hu-Asp2(a), or Hu-Asp2(b) polypeptide comprising culturing the above-described host cell and isolating the relevant polypeptide.

In another aspect, the invention provides isolated Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides, as well as fragments thereof. In a preferred embodiment, the Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides have the amino acid sequence given in SEQ ID NO. 2, SEQ ID NO. 4, or SEQ ID NO.6, respectively. The present invention also describes active forms of Hu-Asp2, methods for preparing such active forms, methods for preparing soluble forms, methods for measuring Hu-Asp2 activity, and substrates for Hu-Asp2 cleavage. The invention also describes antisense oligomers targeting the Hu-Asp1, Hu-Asp2(a) and Hu-Asp2(b) mRNA transcripts and the use of such antisense reagents to decrease such mRNA and consequently the production of the corresponding polypeptide. Isolated antibodies, both polyclonal and monoclonal, that binds specifically to any of the Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides of the invention are also provided.

The invention also provides a method for the identification of an agent that modulates the activity of any of Hu-Asp-1, Hu-Asp2(a), and Hu-Asp2(b). The inventions describes methods to test such agents in cell-free assays to which Hu-Asp2 polypeptide

is added, as well as methods to test such agents in human or other mammalian cells in which Hu-Asp2 is present.

The invention provides for methods for assaying for human Asp1(hu-Asp1)  $\alpha$ -secretase activity comprising contacting the hu-Asp1 protein with an amyloid precursor protein (APP) substrate, wherein the substrate contains an  $\alpha$ -secretase cleavage site; and measuring cleavage of the APP substrate at the  $\alpha$ -secretase cleavage site, thereby assaying hu-Asp1  $\alpha$ -secretase activity. An example of  $\alpha$ -secretase activity is APP processing wherein the APP substrate is cleaved at a site adjacent to the cell membrane (at residues Phe<sup>20</sup>↓Ala<sup>21</sup> in relation to the A $\beta$  peptide). This cleavage results in the release of a soluble, extracellular domain of APP, known as amyloid alpha peptide (sAPP $\alpha$ ), from the cell surface into the cytoplasm. The sAPP $\alpha$  within the cytoplasm can be detected and quantitated thereby measuring  $\alpha$ -secretase activity.

The hu-Asp1 enzyme used in the methods of the invention can be purified and isolated from a cell which is transfected or transformed with a polynucleotide that encodes hu-Asp1, such as SEQ ID NO: 1, or a polynucleotide sequence that encodes the amino acid sequence of SEQ ID NO: 2. Further, the hu-Asp1 protein used in the methods may be a fragment of the amino acid sequence of SEQ ID NO: 2 which retains  $\alpha$ -secretase activity. Possible fragments that may be of use for the methods include those lacking the transmembrane domain amino acids 469-492 of SEQ ID NO: 2, those fragments that lack the cytoplasmic amino acids 493-492 of SEQ ID NO: 2, those fragments that lack the amino terminal amino acids 1-62 of SEQ ID NO: 2 or combinations thereof.

The invention also encompasses methods of assaying for  $\alpha$ -secretase activity where hu-Asp1 protein and its substrate are brought into contact by a growing cell transfected or transformed with a polynucleotide encoding the hu-Asp1 protein or a fragment thereof that retains  $\alpha$ -secretase activity under conditions where the cell expresses hu-Asp1 protein in the presence of the APP substrate. The APP substrate in such circumstances can be exogenously introduced, or more preferably, is expressed by the cell that expresses Asp1. These methods also encompass contacting hu-Asp1 protein with a cell that expresses a polynucleotide that encodes an APP substrate containing an

$\alpha$ -secretase cleavage site. For example, the cell may express a polynucleotide that encodes a polypeptide having an  $\alpha$ -secretase cleavage site comprising the amino acid sequence LVFFAEDF or KLVFFAED. In addition, the APP substrate may comprise any human isoform of APP, such as "normal" APP (APP695), APP 751, or APP770. These APP substrates can be further modified to comprise a carboxy-terminal di-lysine motif.

To measure the cleavage of the substrates for the methods of assaying for  $\alpha$ -secretase activity of the invention, the substrates of the method can be further modified to comprise detectable labels such as radioactive, enzymatic, chemiluminescent or fluorescent labels. In particular, shorter peptide substrates preferably comprise internally quenched labels that result in increased detectability after cleavage of the peptide substrates. The peptide substrates may be modified to have attached a paired fluorophore and quencher including but not limited to 7-amino-4-methyl coumarin and dinitrophenol, respectively, such that cleavage of the peptide by the hu-Asp1 results in increased fluorescence due to physical separation of the fluorophore and quencher. Other paired fluorophores and quenchers include bodipy-tetramethylrhodamine and QSY-5 (Molecular Probes, Inc.) In a variant of this assay, biotin or another suitable tag may be placed on one end of the peptide to anchor the peptide to a substrate assay plate and a fluorophore may be placed at the other end of the peptide. Useful fluorophores include those listed above as well as Europium labels such as W8044 (EG&G Wallac, Inc.). A preferred label is oregon green that may be attached to a Cys residue. Cleavage of the peptide by Asp1 will release the fluorophore or other tag from the plate, allowing compounds to be assayed for inhibition of Asp1 proteolytic cleavage as shown by an increase in retained fluorescence. Preferred colorimetric assays of hu-Asp1 proteolytic activity utilize other suitable substrates that include the P<sub>2</sub> and P<sub>1</sub> amino acids comprising the recognition site for cleavage linked to o-nitrophenol through an amide linkage, such that cleavage by the hu-Asp1 results in an increase in optical density after altering the assay buffer to alkaline pH.

The present invention also provides for methods of assaying for  $\alpha$ -secretase activity comprising contacting hu-Asp1 protein with an APP substrate, determining the level of hu-Asp1  $\alpha$ -secretase activity in the presence and absence of a modulator of hu-

Asp1  $\alpha$ -secretase activity and comparing the hu-Asp1 secretase activity in the presence and absence of the modulator. The modulators determined to increase hu-Asp1  $\alpha$ -secretase activity will be identified as candidate Alzheimer's disease therapeutics. The invention also encompasses methods which comprise a step for treating Alzheimer's disease with identified candidate Alzheimer disease therapeutics. The invention also provides for compositions comprising a candidate Alzheimer's disease therapeutic identified by the  $\alpha$ -secretase assaying methods of the invention. Asp1 modulators that reduce Asp1  $\beta$ -secretase activity and increase Asp1  $\alpha$ -secretase activity are highly preferred. Assays for Asp1  $\beta$ -secretase activity are preferred essentially as described in detail herein for Asp2.

The invention provides for Asp1 protease substrate peptides or fragments thereof, wherein said peptides comprise an amino acid sequence consisting of fifty or fewer amino acids which comprise the Asp1 cleavage site having the amino acid sequence GLALALEP. This peptide was derived from the Asp1 amino acid sequence and the discovery of an apparent Asp1 autocatalytic cleavage in acidic conditions. The Asp1 substrate of the invention may also comprise a detectable label, such as a radioactive label, chemiluminescent label, enzymatic label or a fluorescent label. The fluorescently labeled substrate can consist of internally quenched labels as described above.

The invention also encompasses methods comprising the steps of contacting hu-Asp1 protein with an Asp1 substrate under acidic conditions and determining the level of Asp1 proteolytic activity. An example of Asp1 proteolytic activity is the auto-catalytic processing hu-Asp undergoes in acidic environments, wherein cleavage occurs at an amino acid site surrounding Ala<sup>63</sup> and cleaves the amino terminal amino acids of the hu-Asp1 pro-peptide. The hu-Asp1 pro-peptide refers to a secreted form of Asp1 that has completed intercellular processing which resulted in cleavage of its signal sequence.

For the methods of assaying Asp1 proteolytic activity, the hu-Asp1 may be produced in a cell transformed or transfected with a polynucleotide that encodes hu-Asp1. The hu-Asp1 protein may be isolated and purified from these cells or the method may utilize a cell growing under conditions that it expresses hu-Asp1. The method may also be carried out with a fragment of hu-Asp1 that retains its proteolytic activity. The

fragments provided for by the invention include hu-Asp1 polypeptide sequences which lack the amino acids that encode a transmembrane domain such as amino acids 469-492 of SEQ ID NO: 2 or fragments that lack the cytoplasmic domain such as amino acids 493-518 of SEQ ID NO: 2.

The invention provides for a purified polynucleotide comprising a nucleotide sequence encoding a polypeptide that comprises a fragment of a hu-Asp1 protein, wherein said nucleotide sequence lacks the sequence that encodes amino acids 23-62 of SEQ ID NO: 2, and wherein the polypeptide has hu-Asp1  $\alpha$ -secretase activity. This portion of the Asp1 sequence corresponds to a punitive pro-peptide that is removed, apparently through autocatalysis, under acidic conditions. These polypeptide fragments also include those lacking the amino terminal amino acids 1-62 of SEQ ID NO: 2.

The invention encompasses a purified polynucleotide comprising a nucleotide sequence that hybridizes under stringent conditions to the non-coding strand complementary to SEQ ID NO: 1, wherein the nucleotide sequence encodes a polypeptide having Asp1 proteolytic activity and wherein the polynucleotide lacks nucleotides encoding a transmembrane domain. Further, the invention also provides for a purified polynucleotide sequence comprising a nucleotide sequence that hybridizes under stringent conditions to the non-coding strand complementary to SEQ ID NO: 1, wherein the nucleotide sequence encodes a polypeptide further lacking a pro-peptide domain corresponding to amino acids 23-62 of SEQ ID NO: 2.

The invention also provides for vectors comprising the hu-Asp1 polynucleotide of the invention and host cells transfected or transformed with these vectors. The invention also encompasses host cells transfected or transformed with the hu-Asp1 polynucleotides of the invention.

Another embodiment of the invention is a purified polypeptide comprising a fragment of a hu-Asp1 protein, wherein said polypeptide lacks the hu-Asp1 transmembrane domain and retains hu-Asp1  $\alpha$ -secretase activity. These polypeptides include fragments of hu-Asp1 having the amino acid sequence set forth in SEQ ID NO: 2, and wherein the polypeptide optionally lacks the transmembrane domain amino acids 469-492 of SEQ ID NO: 2, wherein the polypeptide lacks the cytoplasmic domain amino

acids 493-518 of SEQ ID NO: 2 as well. In one variation, the invention provides a polypeptide that lacks amino terminal amino acids 1-62 of SEQ ID NO: 2 but retains Asp1 proteolytic activity. Fragments lacking both the aforementioned amino-terminal and carboxy terminal residues are contemplated.

The invention provides for a polypeptide comprising a fragment of hu-Asp1 having the amino acid sequence set forth in SEQ ID NO: 2 and wherein said polypeptide lacks the amino terminal amino acids 1-62 and retains APP processing activity. For example, referring to the Asp1 sequence in SEQ ID NO: 2, this pre-pro portion would correspond to residues 22 to 62. By performing conventional sequence analysis, the corresponding portions of the Asp1 sequence can also be identified.

Another embodiment of the invention is a polypeptide comprising the amino acid sequence at least 95% identical to a fragment of hu-Asp1 protein, wherein said polypeptide and said fragment lack a transmembrane domain and retain hu-Asp1  $\alpha$ -secretase activity. In addition, the invention embodies a polypeptide comprising an amino acid sequence at least 95% identical to a fragment of hu-Asp1 protein, wherein said polypeptide and said fragment lack the amino terminal amino acids corresponding to the pre-pro portion of hu-Asp1 and retain APP processing activity.

Additional features and variations of the invention will be apparent to those skilled in the art from the entirety of this application, including the drawing and detailed description, and all such features are intended as aspects of the invention. Likewise, features of the invention described herein can be re-combined into additional embodiments that are also intended as aspects of the invention, irrespective of whether the combination of features is specifically mentioned above as an aspect or embodiment of the invention. Also, only such limitations which are described herein as critical to the invention should be viewed as such; variations of the invention lacking limitations which have not been described herein as critical are intended as aspects of the invention.

In addition to the foregoing, the invention includes, as an additional aspect, all embodiments of the invention narrower in scope in any way than the variations specifically mentioned above. Although the applicant(s) invented the full scope of the claims appended hereto, the claims appended hereto are not intended to encompass within



their scope the prior art work of others. Therefore, in the event that statutory prior art within the scope of a claim is brought to the attention of the applicants by a Patent Office or other entity or individual, the applicant(s) reserve the right to exercise amendment rights under applicable patent laws to redefine the subject matter of such a claim to specifically exclude such statutory prior art or obvious variations of statutory prior art from the scope of such a claim. Variations of the invention defined by such amended claims also are intended as aspects of the invention.

### **BRIEF DESCRIPTION OF THE SEQUENCE LISTING**

Sequence ID No. 1: Human Asp-1, nucleotide sequence.

Sequence ID No. 2: Human Asp-1, predicted amino acid sequence.

Sequence ID No. 3: Human Asp-2(a), nucleotide sequence.

Sequence ID No. 4: Human Asp-2(a), predicted amino acid sequence. The Asp2(a) amino acid sequence includes a putative signal peptide comprising residues 1 to 21; and a putative pre-propeptide after the signal peptide that extends through residue 45 (as assessed by processing observed of recombinant Asp2(a) in CHO cells), and a putative propeptide that may extend to at least about residue 57, based on the observation of an observed GRR↓GS sequence which has characteristics of a protease recognition sequence. The Asp2(a) further includes a transmembrane domain comprising residues 455-477, a cytoplasmic domain comprising residues 478-501, and a putative alpha-helical spacer region, comprising residues 420-454, believed to be unnecessary for proteolytic activity, between the protease catalytic domain and the transmembrane domain.

Sequence ID No. 5: Human Asp-2(b), nucleotide sequence.

Sequence ID No. 6: Human Asp-2(b), predicted amino acid sequence. The Asp2(b) amino acid sequence includes a putative signal peptide, pre-propeptide, and propeptide as described above for Asp2(a). The Asp2(b) further includes a transmembrane domain comprising residues 430-452, a cytoplasmic domain comprising residues 453-476, and a putative alpha-helical spacer region, comprising

residues 395-429, believed to be unnecessary for proteolytic activity, between the protease catalytic domain and the transmembrane domain.

Sequence ID No. 7: Murine Asp-2(a), nucleotide sequence.

Sequence ID No. 8: Murine Asp-2(a), predicted amino acid sequence. The proteolytic processing of murine Asp2(a) is believed to be analogous to the processing described above for human Asp2(a). In addition, a variant lacking amino acid residues 190-214 of SEQ ID NO: 8 is specifically contemplated as a murine Asp2(b) polypeptide.

Sequence ID No. 9: Human APP695, nucleotide sequence.

Sequence ID No.10: Human APP695, predicted amino acid sequence.

Sequence ID No.11: Human APP695-Sw, nucleotide sequence.

Sequence ID No.12: Human APP695-Sw, predicted amino acid sequence. In the APP695 isoform, the Sw mutation is characterized by a KM→NL alteration at positions 595-596 (compared to normal APP695).

Sequence ID No.13: Human APP695-VF, nucleotide sequence.

Sequence ID No.14: Human APP695-VF, predicted amino acid sequence. In the APP 695 isoform, the VF mutation is characterized by a V→F alteration at position 642 (compared to normal APP 695).

Sequence ID No.15: Human APP695-KK, nucleotide sequence.

Sequence ID No.16: Human APP695-KK, predicted amino acid sequence.  
(APP695 with two carboxy-terminal lysine residues.)

Sequence ID No.17: Human APP695-Sw-KK, nucleotide sequence.

Sequence ID No.18: Human APP695-Sw-KK, predicted amino acid sequence

Sequence ID No.19: Human APP695-VF-KK, nucleotide sequence

Sequence ID No.20: Human APP695-VF-KK, predicted amino acid sequence

Sequence ID No.21: T7-Human-pro-Asp-2(a)ΔTM, nucleotide sequence

Sequence ID No.22: T7-Human-pro-Asp-2(a)ΔTM, amino acid sequence

Sequence ID No.23: T7-Caspase-Human-pro-Asp-2(a)ΔTM, nucleotide  
sequence

Sequence ID No.24: T7-Caspase-Human-pro-Asp-2(a) $\Delta$ TM, amino acid sequence

Sequence ID No.25: Human-pro-Asp-2(a) $\Delta$ TM (low GC), nucleotide sequence

Sequence ID No.26: Human-pro-Asp-2(a) $\Delta$ TM, (low GC), amino acid sequence

Sequence ID No.27: T7-Caspase-Caspase 8 cleavage-Human-pro-Asp-2(a) $\Delta$ TM, nucleotide sequence

Sequence ID No.28: T7-Caspase-Caspase 8 cleavage-Human-pro-Asp-2(a) $\Delta$ TM, amino acid sequence

Sequence ID No.29: Human Asp-2(a) $\Delta$ TM, nucleotide sequence

Sequence ID No.30: Human Asp-2(a) $\Delta$ TM, amino acid sequence

Sequence ID No.31: Human Asp-2(a) $\Delta$ TM(His)<sub>6</sub>, nucleotide sequence

Sequence ID No. 32: Human Asp-2(a) $\Delta$ TM(His)<sub>6</sub>, amino acid sequence

Sequence ID Nos. 33-49 are short synthetic peptide and oligonucleotide sequences that are described below in the Detailed Description of the Invention.

Sequence ID No. 50: Human Asp2(b) $\Delta$ TM polynucleotide sequence.

Sequence ID No. 51: Human Asp2(b) $\Delta$ TM polypeptide sequence (exemplary variant of Human Asp2(b) lacking transmembrane and intracellular domains of Hu-Asp2(b) set forth in SEQ ID NO: 6.

Sequence ID No. 52: Human Asp2(b) $\Delta$ TM(His)<sub>6</sub> polynucleotide sequence.

Sequence ID No. 53: Human Asp2(b) $\Delta$ TM(His)<sub>6</sub> polypeptide sequence (Human Asp2(b) $\Delta$ TM with six histidine tag attached to C-terminus).

Sequence ID No. 54: Human APP770-encoding polynucleotide sequence.

Sequence ID No. 55: Human APP770 polypeptide sequence. To introduce the KM $\rightarrow$ NL Swedish mutation, residues KM at positions 670-71 are changed to NL. To introduce the V $\rightarrow$ F London mutation, the V residue at position 717 is changed to F.

Sequence ID No. 56: Human APP751 encoding polynucleotide sequence.

Sequence ID No. 57: Human APP751 polypeptide sequence (Human APP751 isoform).

Sequence ID No. 58: Human APP770-KK encoding polynucleotide sequence.

Sequence ID No. 59: Human APP770-KK polypeptide sequence. (Human APP770 isoform to which two C-terminal lysines have been added).

Sequence ID No. 60: Human APP751-KK encoding polynucleotide sequence.

Sequence ID No. 61: Human APP751-KK polypeptide sequence (Human APP751 isoform to which two C-terminal lysines have been added).

Sequence ID Nos. 62-65: Various short peptide sequences described in detail in detailed description.

Sequence ID No. 66: Predicted amino acid sequence of human Asp-1 $\Delta$ TM(His)<sub>6</sub> as described in Example 14.

Sequence ID No. 67: Amino acid sequence of secreted recombinant Asp-1 $\Delta$ TM(His)<sub>6</sub> as described in Example 14.

Sequence ID No. 68: Amino acid sequence of acid-processed form of Asp1 $\Delta$ (His)<sub>6</sub>.

Sequence ID No. 69: Amino acid sequence of the self-activated acid processing site within Asp-1 $\Delta$ TM.

Sequence ID No. 70: Amino acid sequence of a peptide that includes the  $\beta$ -secretase processing site within the Swedish mutant form of APP.

Sequence ID No. 71: Amino acid sequence of a peptide (residues 17-24) that includes the  $\alpha$ -secretase processing site within the A $\beta$  peptide (A $\beta$ <sub>12-28</sub>).

Sequence ID No. 72: Amino acid sequence of a peptide (residues 16-23) that includes the  $\alpha$ -secretase processing site within the A $\beta$  peptide (A $\beta$ <sub>12-28</sub>).

Sequence ID No. 73-74: PCR primers described in Example 14.

Sequence ID No. 75: Amino acid sequence of a  $\gamma$ -secretase substrate polypeptide described in Example 15.

## **BRIEF DESCRIPTION OF THE FIGURES**

Figure 1: Figure 1 shows the nucleotide (SEQ ID NO:1) and predicted amino acid sequence (SEQ ID NO:2) of human Asp1.

**Figure 2:** Figure 2 shows the nucleotide (SEQ ID NO:3) and predicted amino acid sequence (SEQ ID NO:4) of human Asp2(a).

**Figure 3:** Figure 3 shows the nucleotide (SEQ ID NO:5) and predicted amino acid sequence (SEQ ID NO:6) of human Asp2(b). The predicted transmembrane domain of Hu-Asp2(b) is enclosed in brackets.

**Figure 4:** Figure 4 shows the nucleotide (SEQ ID No. 7) and predicted amino acid sequence (SEQ ID No. 8) of murine Asp2(a)

**Figure 5:** Figure 5 shows the BestFit alignment of the predicted amino acid sequences of Hu-Asp2(a) and murine Asp2(a)

**Figure 6:** Figure 6 shows the nucleotide (SEQ ID No. 21) and predicted amino acid sequence (SEQ ID No. 22) of T7-Human-pro-Asp-2(a) $\Delta$ TM

**Figure 7:** Figure 7 shows the nucleotide (SEQ ID No. 23) and predicted amino acid sequence (SEQ ID No. 24) of T7-caspase-Human-pro-Asp-2(a) $\Delta$ TM

**Figure 8:** Figure 8 shows the nucleotide (SEQ ID No. 25) and predicted amino acid sequence (SEQ ID No. 26) of Human-pro-Asp-2(a) $\Delta$ TM (low GC)

**Figure 9:** Western blot showing reduction of CTF99 production by HEK125.3 cells transfected with antisense oligomers targeting the Hu-Asp2 mRNA.

**Figure 10:** Western blot showing increase in CTF99 production in mouse Neuro-2a cells cotransfected with APP-KK with and without Hu-Asp2 only in those cells cotransfected with Hu-Asp2. A further increase in CTF99 production is seen in cells cotransfected with APP-Sw-KK with and without Hu-Asp2 only in those cells cotransfected with Hu-Asp2

**Figure 11:** Figure 11 shows the predicted amino acid sequence (SEQ ID No. 30) of Human-Asp2(a) $\Delta$ TM

**Figure 12:** Figure 11 shows the predicted amino acid sequence (SEQ ID No. 30) of Human-Asp2(a) $\Delta$ TM(His)<sub>6</sub>

## **DETAILED DESCRIPTION OF THE INVENTION**

A few definitions used in this invention follow, most definitions to be used are those that would be used by one ordinarily skilled in the art.

The term "β amyloid peptide" means any peptide resulting from beta secretase cleavage of APP. This includes peptides of 39, 40, 41, 42 and 43 amino acids, extending from the β-secretase cleavage site to 39, 40, 41, 42 and 43 amino acids C-terminal to the β-secretase cleavage site. β amyloid peptide also includes sequences 1-6, SEQ ID NOs. 1-6 of US 5,750,349, issued 12 May 1998 (incorporated into this document by reference).

A β-secretase cleavage fragment disclosed here is called CTF-99, which extends from β-secretase cleavage site to the carboxy terminus of APP.

When an isoform of APP is discussed then what is meant is any APP polypeptide, including APP variants (including mutations), and APP fragments that exists in humans such as those described in US 5,766,846, col 7, lines 45-67, incorporated into this document by reference.

The term "β-amyloid precursor protein" (APP) as used herein is defined as a polypeptide that is encoded by a gene of the same name localized in humans on the long arm of chromosome 21 and that includes "βAP – here "β-amyloid protein" see above, within its carboxyl third. APP is a glycosylated, single-membrane spanning protein expressed in a wide variety of cells in many mammalian tissues. Examples of specific isotypes of APP which are currently known to exist in humans are the 695 amino acid polypeptide described by Kang et. al. (1987) Nature 325:733-736 which is designated as the "normal" APP (SEQ ID NOs: 9-10); the 751 amino acid polypeptide described by Ponte et al. (1988) Nature 331:525-527 (1988) and Tanzi et al. (1988) Nature 331:528-530 (SEQ ID NOs: 56-57); and the 770-amino acid polypeptide described by Kitaguchi et. al. (1988) Nature 331:530-532 (SEQ ID NOs: 54-55). Examples of specific variants of APP include point mutation which can differ in both position and phenotype (for review of known variant mutation see Hardy (1992) Nature Genet. 1:233-234). All references cited here incorporated by reference. The term "APP fragments" as used herein refers to fragments of APP other than those which consist solely of βAP or βAP fragments. That is, APP fragments will include amino acid sequences of APP in addition to those which form intact βAP or a fragment of βAP.

When the term "any amino acid" is used, the amino acids referred to are to be selected from the following, three letter and single letter abbreviations - which may also be used, are provided as follows:

Alanine, Ala, A; Arginine, Arg, R; Asparagine, Asn, N; Aspartic acid, Asp, D; Cysteine, Cys, C; Glutamine, Gln, Q; Glutamic Acid, Glu, E; Glycine, Gly, G; Histidine, His, H; Isoleucine, Ile, I; Leucine, Leu, L; Lysine, Lys, K; Methionine, Met, M; Phenylalanine, Phe, F; Proline, Pro, P; Serine, Ser, S; Threonine, Thr, T; Tryptophan, Trp, W; Tyrosine, Tyr, Y; Valine, Val, V; Aspartic acid or Asparagine, Asx, B; Glutamic acid or Glutamine, Glx, Z; Any amino acid, Xaa, X.

The present invention describes a method to scan gene databases for the simple active site motif characteristic of aspartyl proteases. Eukaryotic aspartyl proteases such as pepsin and renin possess a two-domain structure which folds to bring two aspartyl residues into proximity within the active site. These are embedded in the short tripeptide motif DTG, or more rarely, DSG. Most aspartyl proteases occur as proenzyme whose N-terminus must be cleaved for activation. The DTG or DSG active site motif appears at about residue 65-70 in the proenzyme (prorenin, pepsinogen), but at about residue 25-30 in the active enzyme after cleavage of the N-terminal prodomain. The limited length of the active site motif makes it difficult to search collections of short, expressed sequence tags (EST) for novel aspartyl proteases. EST sequences typically average 250 nucleotides or less, and so would encode 80-90 amino acid residues or less. That would be too short a sequence to span the two active site motifs. The preferred method is to scan databases of hypothetical or assembled protein coding sequences. The present invention describes a computer method to identify candidate aspartyl proteases in protein sequence databases. The method was used to identify seven candidate aspartyl protease sequences in the *Caenorhabditis elegans* genome. These sequences were then used to identify by homology search Hu-Asp1 and two alternative splice variants of Hu-Asp2, designated herein as Hu-Asp2(a) and Hu-Asp2(b).

In a major aspect of the invention disclosed here we provide new information about APP processing. Pathogenic processing of the amyloid precursor protein

(APP) via the A $\beta$  pathway requires the sequential action of two proteases referred to as  $\beta$ -secretase and  $\gamma$ -secretase. Cleavage of APP by the  $\beta$ -secretase and  $\gamma$ -secretase generates the N-terminus and C-terminus of the A $\beta$  peptide, respectively. Because over production of the A $\beta$  peptide, particularly the A $\beta_{1-42}$ , has been implicated in the initiation of Alzheimer's disease, inhibitors of either the  $\beta$ -secretase and/or the  $\gamma$ -secretase have potential in the treatment of Alzheimer's disease. Despite the importance of the  $\beta$ -secretase and  $\gamma$ -secretase in the pathogenic processing of APP, molecular definition of these enzymes has not been accomplished to date. That is, it was not known what enzymes were required for cleavage at either the  $\beta$ -secretase or the  $\gamma$ -secretase cleavage site. The sites themselves were known because APP was known and the A $\beta_{1-42}$  peptide was known, see US 5,766,846 and US 5,837,672, (incorporated by reference, with the exception to reference to "soluble" peptides). But what enzyme was involved in producing the A $\beta_{1-42}$  peptide was unknown.

Alignment of the amino acid sequences of Hu-Asp2 with other known aspartyl proteases reveals a similar domain organization. All of the sequences contain a signal sequence followed by a pro-segment and the catalytic domain containing 2 copies of the aspartyl protease active site motif (DTG/DSG) separated by approximately 180 amino acid residues. Comparison of the processing site for proteolytic removal of the pro-segment in the mature forms of pepsin A, pepsin C, cathepsin D, cathepsin E and renin reveals that the mature forms of these enzymes contain between 31-35 amino acid residues upstream of the first DTG motif. Inspection of this region in the Hu-Asp-2 amino acid sequence indicates a preferred processing site within the sequence GRR↓GS as proteolytic processing of pro-protein precursors commonly occurs at site following dibasic amino acid pairs (*eg.* RR). Also, processing at this site would yield a mature enzyme with 35 amino acid residues upstream of the first DTG, consistent with the processing sites for other aspartyl proteases. In the absence of self-activation of Hu-Asp2 or a knowledge of the endogenous protease that processes Hu-Asp2 at this site, a recombinant form was engineered by introducing a recognition site for the PreSission protease (LEVLFQ↓GP) into the expression plasmids for bacterial, insect cell, and mammalian cell expression of pro-Hu-Asp2. In



each case, the Gly residue in P1' position corresponds to the Gly residue 35 amino acids upstream of the first DTG motif in Hu-Asp2.

The present invention involves the molecular definition of several novel human aspartyl proteases and one of these, referred to as Hu-Asp-2(a) and Hu-Asp2(b), has been characterized in detail. Previous forms of asp1 and asp 2 have been disclosed, see EP 0848062 A2 and EP 0855444A2, inventors David Powel et al., assigned to Smith Kline Beecham Corp. (incorporated by reference). Herein are disclosed old and new forms of Hu-Asp 2. For the first time they are expressed in active form, their substrates are disclosed, and their specificity is disclosed. Prior to this disclosure cell or cell extracts were required to cleave the  $\beta$ -secretase site, now purified protein can be used in assays, also described here. Based on the results of (1) antisense knock out experiments, (2) transient transfection knock in experiments, and (3) biochemical experiments using purified recombinant Hu-Asp-2, we demonstrate that Hu-Asp-2 is the  $\beta$ -secretase involved in the processing of APP. Although the nucleotide and predicted amino acid sequence of Hu-Asp-2(a) has been reported, see above, see EP 0848062 A2 and EP 0855444A2, no functional characterization of the enzyme was disclosed. Here the authors characterize the Hu-Asp-2 enzyme and are able to explain why it is a critical and essential enzyme required in the formation of  $A\beta_{1-42}$ , peptide and possible a critical step in the development of AD.

In another embodiment the present invention also describes a novel splice variant of Hu-Asp2, referred to as Hu-Asp-2(b), that has never before been disclosed.

In another embodiment, the invention provides isolated nucleic acid molecules comprising a polynucleotide encoding a polypeptide selected from the group consisting of human aspartyl protease 1 (Hu-Asp1) and two alternative splice variants of human aspartyl protease-2 (Hu-Asp2), designated herein as Hu-Asp2(a) and Hu-Asp2(b). As used herein, all references to "Hu-Asp2" should be understood to refer to both Hu-Asp2(a) and Hu-Asp2(b). Hu-Asp1 is expressed most abundantly in pancreas and prostate tissues, while Hu-Asp2(a) and Hu-Asp2(b) are expressed most abundantly in pancreas and brain tissues. The invention also provides isolated

Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) polypeptides, as well as fragments thereof which exhibit aspartyl protease activity.

The predicted amino acid sequences of Hu-Asp1, Hu-Asp2(a) and Hu-Asp2(b) share significant homology with previously identified mammalian aspartyl proteases such as pepsinogen A, pepsinogen B, cathepsin D, cathepsin E, and renin. P.B.Szecs, *Scand. J. Clin. Lab. Invest.* 52:(Suppl. 210 5-22 (1992)). These enzymes are characterized by the presence of a duplicated DTG/DSG sequence motif. The Hu-Asp1 and HuAsp2 polypeptides disclosed herein also exhibit extremely high homology with the ProSite consensus motif for aspartyl proteases extracted from the SwissProt database.

The nucleotide sequence given as residues 1-1554 of SEQ ID NO:1 corresponds to the nucleotide sequence encoding Hu-Asp1, the nucleotide sequence given as residues 1-1503 of SEQ ID NO:3 corresponds to the nucleotide sequence encoding Hu-Asp2(a), and the nucleotide sequence given as residues 1-1428 of SEQ ID NO:5 corresponds to the nucleotide sequence encoding Hu-Asp2(b). The isolation and sequencing of DNA encoding Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b) is described below in Examples 1 and 2.

As is described in Examples 1 and 2, automated sequencing methods were used to obtain the nucleotide sequence of Hu-Asp1, Hu-Asp2(a), and Hu-Asp-2(b). The Hu-Asp nucleotide sequences of the present invention were obtained for both DNA strands, and are believed to be 100% accurate. However, as is known in the art, nucleotide sequence obtained by such automated methods may contain some errors. Nucleotide sequences determined by automation are typically at least about 90%, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of a given nucleic acid molecule. The actual sequence may be more precisely determined using manual sequencing methods, which are well known in the art. An error in sequence which results in an insertion or deletion of one or more nucleotides may result in a frame shift in translation such that the predicted amino acid sequence will differ from that which would be predicted from the actual nucleotide sequence of the nucleic acid molecule, starting at the point of the mutation.

The Hu-Asp DNA of the present invention includes cDNA, chemically synthesized DNA, DNA isolated by PCR, genomic DNA, and combinations thereof. Genomic Hu-Asp DNA may be obtained by screening a genomic library with the Hu-Asp2 cDNA described herein, using methods that are well known in the art, or with oligonucleotides chosen from the Hu-Asp2 sequence that will prime the polymerase chain reaction (PCR). RNA transcribed from Hu-Asp DNA is also encompassed by the present invention.

Due to the degeneracy of the genetic code, two DNA sequences may differ and yet encode identical amino acid sequences. The present invention thus provides isolated nucleic acid molecules having a polynucleotide sequence encoding any of the Hu-Asp polypeptides of the invention, wherein said polynucleotide sequence encodes a Hu-Asp polypeptide having the complete amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, or fragments thereof.

Also provided herein are purified Hu-Asp polypeptides, both recombinant and non-recombinant. Most importantly, methods to produce Hu-Asp2 polypeptides in active form are provided. These include production of Hu-Asp2 polypeptides and variants thereof in bacterial cells, insect cells, and mammalian cells, also in forms that allow secretion of the Hu-Asp2 polypeptide from bacterial, insect or mammalian cells into the culture medium, also methods to produce variants of Hu-Asp2 polypeptide incorporating amino acid tags that facilitate subsequent purification. In a preferred embodiment of the invention the Hu-Asp2 polypeptide is converted to a proteolytically active form either in transformed cells or after purification and cleavage by a second protease in a cell-free system, such active forms of the Hu-Asp2 polypeptide beginning with the N-terminal sequence TQHGIR or ETDEEP. The sequence TQHGIR represents the amino-terminus of Asp2(a) or Asp2(b) beginning with residue 22 of SEQ ID NO: 4 or 6, after cleavage of a putative 21 residue signal peptide. Recombinant Asp2(a) expressed in and purified from insect cells was observed to have this amino terminus, presumably as a result of cleavage by a signal peptidase. The sequence ETDEEP represents the amino-terminus of Asp2(a) or Asp2(b) beginning with residue 46 of SEQ ID NO: 4 or 6, as observed when Asp2(a)

has been recombinantly produced in CHO cells (presumably after cleavage by both a rodent signal peptidase and another rodent peptidase that removes a propeptide sequence). The Asp2(a) produced in the CHO cells possesses  $\beta$ -secretase activity, as described in greater detail in Examples 11 and 12. Variants and derivatives, including fragments, of Hu-Asp proteins having the native amino acid sequences given in SEQ ID Nos: 2, 4, and 6 that retain any of the biological activities of Hu-Asp are also within the scope of the present invention. Of course, one of ordinary skill in the art will readily be able to determine whether a variant, derivative, or fragment of a Hu-Asp protein displays Hu-Asp activity by subjecting the variant, derivative, or fragment to a standard aspartyl protease assay. Fragments of Hu-Asp within the scope of this invention include those that contain the active site domain containing the amino acid sequence DTG, fragments that contain the active site domain amino acid sequence DSG, fragments containing both the DTG and DSG active site sequences, fragments in which the spacing of the DTG and DSG active site sequences has been lengthened, fragments in which the spacing has been shortened. Also within the scope of the invention are fragments of Hu-Asp in which the transmembrane domain has been removed to allow production of Hu-Asp2 in a soluble form. In another embodiment of the invention, the two halves of Hu-Asp2, each containing a single active site DTG or DSG sequence can be produced independently as recombinant polypeptides, then combined in solution where they reconstitute an active protease.

Thus, the invention provides a purified polypeptide comprising a fragment of a mammalian Asp2 protein, wherein said fragment lacks the Asp2 transmembrane domain of said Asp2 protein, and wherein the polypeptide and the fragment retain the  $\beta$ -secretase activity of said mammalian Asp2 protein. In a preferred embodiment, the purified polypeptide comprises a fragment of a human Asp2 protein that retains the  $\beta$ -secretase activity of the human Asp2 protein from which it was derived. Examples include:

a purified polypeptide that comprises a fragment of Asp2(a) having the amino acid sequence set forth in SEQ ID NO: 4, wherein the polypeptide lacks transmembrane domain amino acids 455 to 477 of SEQ ID NO: 4;

a purified polypeptide as described in the preceding paragraph that further lacks cytoplasmic domain amino acids 478 to 501 of SEQ ID NO: 4;

a purified polypeptide as described in either of the preceding paragraphs that further lacks amino acids 420-454 of SEQ ID NO: 4, which constitute a putative alpha helical region between the catalytic domain and the transmembrane domain that is believed to be unnecessary for  $\beta$ -secretase activity;

a purified polypeptide that comprises an amino acid sequence that includes amino acids 58 to 419 of SEQ ID NO: 4, and that lacks amino acids 22 to 57 of SEQ ID NO: 4;

a purified polypeptide that comprises an amino acid sequence that includes amino acids 46 to 419 of SEQ ID NO: 4, and that lacks amino acids 22 to 45 of SEQ ID NO: 4;

a purified polypeptide that comprises an amino acid sequence that includes amino acids 22 to 454 of SEQ ID NO: 4.

a purified polypeptide that comprises a fragment of Asp2(b) having the amino acid sequence set forth in SEQ ID NO: 6, and wherein said polypeptide lacks transmembrane domain amino acids 430 to 452 of SEQ ID NO: 6;

a purified polypeptide as described in the preceding paragraph that further lacks cytoplasmic domain amino acids 453 to 476 of SEQ ID NO: 6;

a purified polypeptide as described in either of the preceding two paragraphs that further lacks amino acids 395-429 of SEQ ID NO: 4, which constitute a putative alpha helical region between the catalytic domain and the transmembrane domain that is believed to be unnecessary for  $\beta$ -secretase activity;

a purified polypeptide comprising an amino acid sequence that includes amino acids 58 to 394 of SEQ ID NO: 4, and that lacks amino acids 22 to 57 of SEQ ID NO: 4;

a purified polypeptide comprising an amino acid sequence that includes amino acids 46 to 394 of SEQ ID NO: 4, and that lacks amino acids 22 to 45 of SEQ ID NO: 4; and

a purified polypeptide comprising an amino acid sequence that includes amino acids 22 to 429 of SEQ ID NO: 4.

Also included as part of the invention is a purified polynucleotide comprising a nucleotide sequence that encodes such polypeptides; a vector comprising a polynucleotide that encodes such polypeptides; and a host cell transformed or transfected with such a polynucleotide or vector.

Hu-Asp variants may be obtained by mutation of native Hu-Asp-encoding nucleotide sequences, for example. A Hu-Asp variant, as referred to herein, is a polypeptide substantially homologous to a native Hu-Asp polypeptide but which has an amino acid sequence different from that of native Hu-Asp because of one or more deletions, insertions, or substitutions in the amino acid sequence. The variant amino acid or nucleotide sequence is preferably at least about 80% identical, more preferably at least about 90% identical, and most preferably at least about 95% identical, to a native Hu-Asp sequence. Thus, a variant nucleotide sequence which contains, for example, 5 point mutations for every one hundred nucleotides, as compared to a native Hu-Asp gene, will be 95% identical to the native protein. The percentage of sequence identity, also termed homology, between a native and a variant Hu-Asp sequence may also be determined, for example, by comparing the two sequences using any of the computer programs commonly employed for this purpose, such as the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison Wisconsin), which uses the algorithm of Smith and Waterman (*Adv. Appl. Math.* 2: 482-489 (1981)).

Alterations of the native amino acid sequence may be accomplished by any of a number of known techniques. For example, mutations may be introduced at particular locations by procedures well known to the skilled artisan, such as oligonucleotide-directed mutagenesis, which is described by Walder *et al.* (*Gene* 42:133 (1986)); Bauer *et al.* (*Gene* 37:73 (1985)); Craik (*BioTechniques*, January

1985, pp. 12-19); Smith *et al.* (*Genetic Engineering: Principles and Methods*, Plenum Press (1981)); and U.S. Patent Nos. 4,518,584 and 4,737,462.

Hu-Asp variants within the scope of the invention may comprise conservatively substituted sequences, meaning that one or more amino acid residues of a Hu-Asp polypeptide are replaced by different residues that do not alter the secondary and/or tertiary structure of the Hu-Asp polypeptide. Such substitutions may include the replacement of an amino acid by a residue having similar physicochemical properties, such as substituting one aliphatic residue (Ile, Val, Leu or Ala) for another, or substitution between basic residues Lys and Arg, acidic residues Glu and Asp, amide residues Gln and Asn, hydroxyl residues Ser and Tyr, or aromatic residues Phe and Tyr. Further information regarding making phenotypically silent amino acid exchanges may be found in Bowie *et al.*, *Science* 247:1306-1310 (1990). Other Hu-Asp variants which might retain substantially the biological activities of Hu-Asp are those where amino acid substitutions have been made in areas outside functional regions of the protein.

In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent conditions to a portion of the nucleic acid molecules described above, *e.g.*, to at least about 15 nucleotides, preferably to at least about 20 nucleotides, more preferably to at least about 30 nucleotides, and still more preferably to at least about from 30 to at least about 100 nucleotides, of one of the previously described nucleic acid molecules. Such portions of nucleic acid molecules having the described lengths refer to, *e.g.*, at least about 15 contiguous nucleotides of the reference nucleic acid molecule. By stringent hybridization conditions is intended overnight incubation at about 42°C for about 2.5 hours in 6 X SSC/0.1% SDS, followed by washing of the filters four times for 15 minutes in 1.0 X SSC at 65°C, 0.1% SDS.

Fragments of the Hu-Asp encoding nucleic acid molecules described herein, as well as polynucleotides capable of hybridizing to such nucleic acid molecules may be used as a probe or as primers in a polymerase chain reaction (PCR). Such probes may be used, *e.g.*, to detect the presence of Hu-Asp nucleic acids in *in vitro* assays, as well

as in Southern and northern blots. Cell types expressing Hu-Asp may also be identified by the use of such probes. Such procedures are well known, and the skilled artisan will be able to choose a probe of a length suitable to the particular application. For PCR, 5' and 3' primers corresponding to the termini of a desired Hu-Asp nucleic acid molecule are employed to isolate and amplify that sequence using conventional techniques.

Other useful fragments of the Hu-Asp nucleic acid molecules are antisense or sense oligonucleotides comprising a single stranded nucleic acid sequence capable of binding to a target Hu-Asp mRNA (using a sense strand), or Hu-Asp DNA (using an antisense strand) sequence. In a preferred embodiment of the invention these Hu-Asp antisense oligonucleotides reduce Hu-Asp mRNA and consequent production of Hu-Asp polypeptides.

In another aspect, the invention includes Hu-Asp polypeptides with or without associated native pattern glycosylation. Both Hu-Asp1 and Hu-Asp2 have canonical acceptor sites for Asn-linked sugars, with Hu-Asp1 having two of such sites, and Hu-Asp2 having four. Hu-Asp expressed in yeast or mammalian expression systems (discussed below) may be similar to or significantly different from a native Hu-Asp polypeptide in molecular weight and glycosylation pattern. Expression of Hu-Asp in bacterial expression systems will provide non-glycosylated Hu-Asp.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. Hu-Asp polypeptides may be recovered and purified from tissues, cultured cells, or recombinant cell cultures by well-known methods, including ammonium sulfate or ethanol precipitation, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, lectin chromatography, and high performance liquid chromatography (HPLC). In a preferred embodiment, an amino acid tag is added to the Hu-Asp polypeptide using genetic engineering techniques that are well known to practitioners of the art which include addition of six histidine amino acid residues to allow purification by binding to nickel immobilized on a suitable support, epitopes for



polyclonal or monoclonal antibodies including but not limited to the T7 epitope, the myc epitope, and the V5a epitope, and fusion of Hu-Asp2 to suitable protein partners including but not limited to glutathione-S-transferase or maltose binding protein. In a preferred embodiment these additional amino acid sequences are added to the C-terminus of Hu-Asp but may be added to the N-terminus or at intervening positions within the Hu-Asp2 polypeptide.

The present invention also relates to vectors comprising the polynucleotide molecules of the invention, as well as host cell transformed with such vectors. Any of the polynucleotide molecules of the invention may be joined to a vector, which generally includes a selectable marker and an origin of replication, for propagation in a host. Because the invention also provides Hu-Asp polypeptides expressed from the polynucleotide molecules described above, vectors for the expression of Hu-Asp are preferred. The vectors include DNA encoding any of the Hu-Asp polypeptides described above or below, operably linked to suitable transcriptional or translational regulatory sequences, such as those derived from a mammalian, microbial, viral, or insect gene. Examples of regulatory sequences include transcriptional promoters, operators, or enhancers, mRNA ribosomal binding sites, and appropriate sequences which control transcription and translation. Nucleotide sequences are operably linked when the regulatory sequence functionally relates to the DNA encoding Hu-Asp. Thus, a promoter nucleotide sequence is operably linked to a Hu-Asp DNA sequence if the promoter nucleotide sequence directs the transcription of the Hu-Asp sequence.

Selection of suitable vectors to be used for the cloning of polynucleotide molecules encoding Hu-Asp, or for the expression of Hu-Asp polypeptides, will of course depend upon the host cell in which the vector will be transformed, and, where applicable, the host cell from which the Hu-Asp polypeptide is to be expressed. Suitable host cells for expression of Hu-Asp polypeptides include prokaryotes, yeast, and higher eukaryotic cells, each of which is discussed below.

The Hu-Asp polypeptides to be expressed in such host cells may also be fusion proteins which include regions from heterologous proteins. Such regions may be included to allow, *e.g.*, secretion, improved stability, or facilitated purification of the

polypeptide. For example, a sequence encoding an appropriate signal peptide can be incorporated into expression vectors. A DNA sequence for a signal peptide (secretory leader) may be fused inframe to the Hu-Asp sequence so that Hu-Asp is translated as a fusion protein comprising the signal peptide. A signal peptide that is functional in the intended host cell promotes extracellular secretion of the Hu-Asp polypeptide. Preferably, the signal sequence will be cleaved from the Hu-Asp polypeptide upon secretion of Hu-Asp from the cell. Nonlimiting examples of signal sequences that can be used in practicing the invention include the yeast Ifactor and the honeybee melatin leader in sf9 insect cells.

In a preferred embodiment, the Hu-Asp polypeptide will be a fusion protein which includes a heterologous region used to facilitate purification of the polypeptide. Many of the available peptides used for such a function allow selective binding of the fusion protein to a binding partner. For example, the Hu-Asp polypeptide may be modified to comprise a peptide to form a fusion protein which specifically binds to a binding partner, or peptide tag. Nonlimiting examples of such peptide tags include the 6-His tag, thioredoxin tag, hemagglutinin tag, GST tag, and OmpA signal sequence tag. As will be understood by one of skill in the art, the binding partner which recognizes and binds to the peptide may be any molecule or compound including metal ions (*e.g.*, metal affinity columns), antibodies, or fragments thereof, and any protein or peptide which binds the peptide, such as the FLAG tag.

Suitable host cells for expression of Hu-Asp polypeptides includes prokaryotes, yeast, and higher eukaryotic cells. Suitable prokaryotic hosts to be used for the expression of Hu-Asp include bacteria of the genera *Escherichia*, *Bacillus*, and *Salmonella*, as well as members of the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*. For expression in, *e.g.*, *E. coli*, a Hu-Asp polypeptide may include an N-terminal methionine residue to facilitate expression of the recombinant polypeptide in a prokaryotic host. The N-terminal Met may optionally then be cleaved from the expressed Hu-Asp polypeptide. Other N-terminal amino acid residues can be added to the Hu-Asp polypeptide to facilitate expression in *Escherichia coli* including but not limited to the T7 leader sequence, the T7-caspase 8 leader sequence, as well as

others leaders including tags for purification such as the 6-His tag (Example 9). Hu-Asp polypeptides expressed in *E. coli* may be shortened by removal of the cytoplasmic tail, the transmembrane domain, or the membrane proximal region. Hu-Asp polypeptides expressed in *E. coli* may be obtained in either a soluble form or as an insoluble form which may or may not be present as an inclusion body. The insoluble polypeptide may be rendered soluble by guanidine HCl, urea or other protein denaturants, then refolded into a soluble form before or after purification by dilution or dialysis into a suitable aqueous buffer. If the inactive proform of the Hu-Asp was produced using recombinant methods, it may be rendered active by cleaving off the prosegment with a second suitable protease such as human immunodeficiency virus protease.

Expression vectors for use in prokaryotic hosts generally comprises one or more phenotypic selectable marker genes. Such genes generally encode, *e.g.*, a protein that confers antibiotic resistance or that supplies an auxotrophic requirement. A wide variety of such vectors are readily available from commercial sources. Examples include pSPORT vectors, pGEM vectors (Promega), pPROEX vectors (LTI, Bethesda, MD), Bluescript vectors (Stratagene), pET vectors (Novagen) and pQE vectors (Qiagen).

Hu-Asp may also be expressed in yeast host cells from genera including *Saccharomyces*, *Pichia*, and *Kluveromyces*. Preferred yeast hosts are *S. cerevisiae* and *P. pastoris*. Yeast vectors will often contain an origin of replication sequence from a 2T yeast plasmid, an autonomously replicating sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination, and a selectable marker gene. Vectors replicable in both yeast and *E. coli* (termed shuttle vectors) may also be used. In addition to the above-mentioned features of yeast vectors, a shuttle vector will also include sequences for replication and selection in *E. coli*. Direct secretion of Hu-Asp polypeptides expressed in yeast hosts may be accomplished by the inclusion of nucleotide sequence encoding the yeast I-factor leader sequence at the 5' end of the Hu-Asp-encoding nucleotide sequence.

Insect host cell culture systems may also be used for the expression of Hu-Asp polypeptides. In a preferred embodiment, the Hu-Asp polypeptides of the invention are expressed using an insect cell expression system (*see* Example 10). Additionally, a baculovirus expression system can be used for expression in insect cells as reviewed by Luckow and Summers, *Bio/Technology* 6:47 (1988).

In another preferred embodiment, the Hu-Asp polypeptide is expressed in mammalian host cells. Nonlimiting examples of suitable mammalian cell lines include the COS7 line of monkey kidney cells (Gluzman *et al.*, *Cell* 23:175 (1981)), human embryonic kidney cell line 293, and Chinese hamster ovary (CHO) cells. Preferably, Chinese hamster ovary (CHO) cells are used for expression of Hu-Asp proteins (Example 11).

The choice of a suitable expression vector for expression of the Hu-Asp polypeptides of the invention will of course depend upon the specific mammalian host cell to be used, and is within the skill of the ordinary artisan. Examples of suitable expression vectors include pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). A preferred vector for expression of Hu-Asp polypeptides is pcDNA3.1-Hygro (Invitrogen). Expression vectors for use in mammalian host cells may include transcriptional and translational control sequences derived from viral genomes. Commonly used promoter sequences and enhancer sequences which may be used in the present invention include, but are not limited to, those derived from human cytomegalovirus (CMV), Adenovirus 2, Polyoma virus, and Simian virus 40 (SV40). Methods for the construction of mammalian expression vectors are disclosed, for example, in Okayama and Berg (*Mol. Cell. Biol.* 3:280 (1983)); Cosman *et al.* (*Mol. Immunol.* 23:935 (1986)); Cosman *et al.* (*Nature* 312:768 (1984)); EP-A-0367566; and WO 91/18982.

The polypeptides of the present invention may also be used to raise polyclonal and monoclonal antibodies, which are useful in diagnostic assays for detecting Hu-Asp polypeptide expression. Such antibodies may be prepared by conventional techniques. See, for example, *Antibodies: A Laboratory Manual*, Harlow and Land (eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (1988);

*Monoclonal Antibodies, Hybridomas: A New Dimension in Biological Analyses*, Kennet *et al.* (eds.), Plenum Press, New York (1980). Synthetic peptides comprising portions of Hu-Asp containing 5 to 20 amino acids may also be used for the production of polyclonal or monoclonal antibodies after linkage to a suitable carrier protein including but not limited to keyhole limpet hemacyanin (KLH), chicken ovalbumin, or bovine serum albumin using various cross-linking reagents including carbodimides, glutaraldehyde, or if the peptide contains a cysteine, N-methylmaleimide. A preferred peptide for immunization when conjugated to KLH contains the C-terminus of Hu-Asp1 or Hu-Asp2 comprising QRRPRDPEVVNDESSLVRHRWK (SEQ ID NO: 2, residues 497-518) or LRQQHDDFADDISLLK (SEQ ID NO:4, residues 486-501), respectively. See SEQ ID Nos. 33-34.

The Hu-Asp nucleic acid molecules of the present invention are also valuable for chromosome identification, as they can hybridize with a specific location on a human chromosome. Hu-Asp1 has been localized to chromosome 21, while Hu-Asp2 has been localized to chromosome 11q23.3-24.1. There is a current need for identifying particular sites on the chromosome, as few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. The relationship between genes and diseases that have been mapped to the same chromosomal region can then be identified through linkage analysis, wherein the coinheritance of physically adjacent genes is determined. Whether a gene appearing to be related to a particular disease is in fact the cause of the disease can then be determined by comparing the nucleic acid sequence between affected and unaffected individuals.

In another embodiment, the invention relates to a method of assaying Hu-Asp function, specifically Hu-Asp2 function which involves incubating in solution the Hu-Asp polypeptide with a suitable substrate including but not limited to a synthetic peptide containing the  $\beta$ -secretase cleavage site of APP, preferably one containing the

mutation found in a Swedish kindred with inherited AD in which KM is changed to NL, such peptide comprising the sequence SEVNLDAEFR in an acidic buffering solution, preferably an acidic buffering solution of pH5.5 (see Example 12) using cleavage of the peptide monitored by high performance liquid chromatography as a measure of Hu-Asp proteolytic activity. Preferred assays for proteolytic activity utilize internally quenched peptide assay substrates. Such suitable substrates include peptides which have attached a paired fluorphore and quencher including but not limited to 7-amino-4-methyl coumarin and dinitrophenol, respectively, such that cleavage of the peptide by the Hu-Asp results in increased fluorescence due to physical separation of the fluorphore and quencher. Other paired fluorphores and quenchers include bodipy-tetramethylrhodamine and QSY-5 (Molecular Probes, Inc.). In a variant of this assay, biotin or another suitable tag may be placed on one end of the peptide to anchor the peptide to a substrate assay plate and a fluorphore may be placed at the other end of the peptide. Useful fluorphores include those listed above as well as Europium labels such as W8044 (EG&g Wallac, Inc.). Cleavage of the peptide by Asp2 will release the fluorphore or other tag from the plate, allowing compounds to be assayed for inhibition of Asp2 proteolytic cleavage as shown by an increase in retained fluorescence. Preferred colorimetric assays of Hu-Asp proteolytic activity utilize other suitable substrates that include the P2 and P1 amino acids comprising the recognition site for cleavage linked to o-nitrophenol through an amide linkage, such that cleavage by the Hu-Asp results in an increase in optical density after altering the assay buffer to alkaline pH.

In another embodiment, the invention relates to a method for the identification of an agent that increases the activity of a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), the method comprising

- (a) determining the activity of said Hu-Asp polypeptide in the presence of a test agent and in the absence of a test agent; and
- (b) comparing the activity of said Hu-Asp polypeptide determined in the presence of said test agent to the activity of said Hu-Asp polypeptide determined in the absence of said test agent;

whereby a higher level of activity in the presence of said test agent than in the absence of said test agent indicates that said test agent has increased the activity of said Hu-Asp polypeptide. Such tests can be performed with Hu-Asp polypeptide in a cell free system and with cultured cells that express Hu-Asp as well as variants or isoforms thereof.

In another embodiment, the invention relates to a method for the identification of an agent that decreases the activity of a Hu-Asp polypeptide selected from the group consisting of Hu-Asp1, Hu-Asp2(a), and Hu-Asp2(b), the method comprising

(a) determining the activity of said Hu-Asp polypeptide in the presence of a test agent and in the absence of a test agent; and

(b) comparing the activity of said Hu-Asp polypeptide determined in the presence of said test agent to the activity of said Hu-Asp polypeptide determined in the absence of said test agent; whereby a lower level of activity in the presence of said test agent than in the absence of said test agent indicates that said test agent has decreased the activity of said Hu-Asp polypeptide. Such tests can be performed with Hu-Asp polypeptide in a cell free system and with cultured cells that express Hu-Asp as well as variants or isoforms thereof.

In another embodiment, the invention relates to a novel cell line (HEK125.3 cells) for measuring processing of amyloid  $\beta$  peptide ( $A\beta$ ) from the amyloid protein precursor (APP). The cells are stable transformants of human embryonic kidney 293 cells (HEK293) with a bicistronic vector derived from pIRES-EGFP (Clontech) containing a modified human APP cDNA, an internal ribosome entry site and an enhanced green fluorescent protein (EGFP) cDNA in the second cistron. The APP cDNA was modified by adding two lysine codons to the carboxyl terminus of the APP coding sequence. This increases processing of  $A\beta$  peptide from human APP by 2-4 fold. This level of  $A\beta$  peptide processing is 60 fold higher than is seen in nontransformed HEK293 cells. HEK125.3 cells will be useful for assays of compounds that inhibit  $A\beta$  peptide processing. This invention also includes addition of two lysine residues to the C-terminus of other APP isoforms including the 751 and 770 amino acid isoforms, to isoforms of APP having mutations found in human AD

including the Swedish KM→NL and V717→F mutations, to C-terminal fragments of APP, such as those beginning with the β-secretase cleavage site, to C-terminal fragments of APP containing the β-secretase cleavage site which have been operably linked to an N-terminal signal peptide for membrane insertion and secretion, and to C-terminal fragments of APP which have been operably linked to an N-terminal signal peptide for membrane insertion and secretion and a reporter sequence including but not limited to green fluorescent protein or alkaline phosphatase, such that β-secretase cleavage releases the reporter protein from the surface of cells expressing the polypeptide.

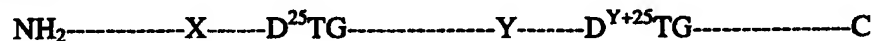
Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

### **Example 1**

#### **Development of a Search Algorithm Useful for the Identification of Aspartyl Proteases, and Identification of *C. elegans* Aspartyl Protease Genes in Wormpep 12**

##### ***Materials and Methods:***

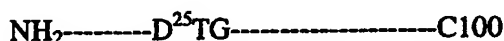
Classical aspartyl proteases such as pepsin and renin possess a two-domain structure which folds to bring two aspartyl residues into proximity within the active site. These are embedded in the short tripeptide motif DTG, or more rarely, DSG. The DTG or DSG active site motif appears at about residue 25-30 in the enzyme, but at about 65-70 in the proenzyme (prorenin, pepsinogen). This motif appears again about 150-200 residues downstream. The proenzyme is activated by cleavage of the N-terminal prodomain. This pattern exemplifies the double domain structure of the modern day aspartyl enzymes which apparently arose by gene duplication and divergence. Thus;



where X denotes the beginning of the enzyme, following the N-terminal prodomain, and Y denotes the center of the molecule where the gene repeat begins again.



In the case of the retroviral enzymes such as the HIV protease, they represent only a half of the two-domain structures of well-known enzymes like pepsin, cathepsin D, renin, etc. They have no prosegment, but are carved out of a polyprotein precursor containing the *gag* and *pol* proteins of the virus. They can be represented by:



This "monomer" only has about 100 aa, so is extremely parsimonious as compared to the other aspartyl protease "dimers" which have of the order of 330 or so aa, not counting the N-terminal prodomain.

The limited length of the eukaryotic aspartyl protease active site motif makes it difficult to search EST collections for novel sequences. EST sequences typically average 250 nucleotides, and so in this case would be unlikely to span both aspartyl protease active site motifs. Instead, we turned to the *C. elegans* genome. The *C. elegans* genome is estimated to contain around 13,000 genes. Of these, roughly 12,000 have been sequenced and the corresponding hypothetical open reading frame (ORF) has been placed in the database Wormpep12. We used this database as the basis for a whole genome scan of a higher eukaryote for novel aspartyl proteases, using an algorithm that we developed specifically for this purpose. The following AWK script for locating proteins containing two DTG or DSG motifs was used for the search, which was repeated four times to recover all pairwise combinations of the aspartyl motif.

```
BEGIN{RS=">"}          /* defines ">" as record separator for FASTA format */
{
  pos = index($0,"DTG")  /* finds "DTG" in record */
  if (pos>0) {
    rest = substr($0,pos+3) /*get rest of record after first DTG*/
    pos2 = index(rest,"DTG") /*find second DTG*/
    if (pos2>0) printf ("%s%s\n",>",$0)} /*report hits*/
  }
}
```

The AWK script shown above was used to search Wormpep12, which was downloaded from <ftp.sanger.ac.uk/pub/databases/wormpep>, for sequence entries

containing at least two DTG or DSG motifs. Using AWK limited each record to 3000 characters or less. Thus, 35 or so larger records were eliminated manually from Wormpep12 as in any case these were unlikely to encode aspartyl proteases.

#### ***Results and Discussion:***

The Wormpep 12 database contains 12,178 entries, although some of these (<10%) represent alternatively spliced transcripts from the same gene. Estimates of the number of genes encoded in the *C. elegans* genome is on the order of 13,000 genes, so Wormpep12 may be estimated to cover greater than 90% of the *C. elegans* genome.

Eukaryotic aspartyl proteases contain a two-domain structure, probably arising from ancestral gene duplication. Each domain contains the active site motif D(S/T)G located from 20-25 amino acid residues into each domain. The retroviral (e.g., HIV protease) or retrotransposon proteases are homodimers of subunits which are homologous to a single eukaryotic aspartyl protease domain. An AWK script was used to search the Wormpep12 database for proteins in which the D(S/T)G motif occurred at least twice. This identified >60 proteins with two DTG or DSG motifs. Visual inspection was used to select proteins in which the position of the aspartyl domains was suggestive of a two-domain structure meeting the criteria described above.

In addition, the PROSITE eukaryotic and viral aspartyl protease active site pattern PS00141 was used to search Wormpep12 for candidate aspartyl proteases. (Bairoch A., Bucher P., Hofmann K., The PROSITE database: its status in 1997, *Nucleic Acids Res.* 24:217-221(1997)). This generated an overlapping set of Wormpep12 sequences. Of these, seven sequences contained two DTG or DSG motifs and the PROSITE aspartyl protease active site pattern. Of these seven, three were found in the same cosmid clone (F21F8.3, F21F8.4, and F21F8.7) suggesting that they represent a family of proteins that arose by ancestral gene duplication. Two other ORFs with extensive homology to F21F8.3, F21F8.4 and F21F8.7 are present in the same gene cluster (F21F8.2 and F21F8.6), however, these contain only a single DTG motif. Exhaustive BLAST searches with these seven sequences against

Wormpep12 failed to reveal additional candidate aspartyl proteases in the *C. elegans* genome containing two repeats of the DTG or DSG motif.

BLASTX search with each *C. elegans* sequence against SWISS-PROT, GenPep and TREMBL revealed that R12H7.2 was the closest worm homologue to the known mammalian aspartyl proteases, and that T18H9.2 was somewhat more distantly related, while CEASP1, F21F8.3, F21F8.4, and F21F8.7 formed a subcluster which had the least sequence homology to the mammalian sequences.

#### ***Discussion:***

APP, the presenilins, and p35, the activator of cdk5, all undergo intracellular proteolytic processing at sites which conform to the substrate specificity of the HIV protease. Dysregulation of a cellular aspartyl protease with the same substrate specificity, might therefore provide a unifying mechanism for causation of the plaque and tangle pathologies in AD. Therefore, we sought to identify novel human aspartyl proteases. A whole genome scan in *C. elegans* identified seven open reading frames that adhere to the aspartyl protease profile that we had identified. These seven aspartyl proteases probably comprise the complete complement of such proteases in a simple, multicellular eukaryote. These include four closely related aspartyl proteases unique to *C. elegans* which probably arose by duplication of an ancestral gene. The other three candidate aspartyl proteases (T18H9.2, R12H7.2 and C11D2.2) were found to have homology to mammalian gene sequences.

#### **Example 2**

##### **Identification of Novel Human Aspartyl Proteases Using Database Mining by Genome Bridging**

#### ***Materials and Methods:***

Computer-assisted analysis of EST databases, cDNA , and predicted polypeptide sequences:

Exhaustive homology searches of EST databases with the CEASP1, F21F8.3, F21F8.4, and F21F8.7 sequences failed to reveal any novel mammalian homologues. TBLASTN searches with R12H7.2 showed homology to cathepsin D, cathepsin E, pepsinogen A, pepsinogen C and renin, particularly around the DTG motif within the

active site, but also failed to identify any additional novel mammalian aspartyl proteases. This indicates that the *C. elegans* genome probably contains only a single lysosomal aspartyl protease which in mammals is represented by a gene family that arose through duplication and consequent modification of an ancestral gene.

TBLASTN searches with T18H9.2, the remaining *C. elegans* sequence, identified several ESTs which assembled into a contig encoding a novel human aspartyl protease (Hu-ASP1). As is described above in Example 1, BLASTX search with the Hu-ASP1 contig against SWISS-PROT revealed that the active site motifs in the sequence aligned with the active sites of other aspartyl proteases. Exhaustive, repetitive rounds of BLASTN searches against LifeSeq, LifeSeqFL, and the public EST collections identified 102 EST from multiple cDNA libraries that assembled into a single contig. The 51 sequences in this contig found in public EST collections also have been assembled into a single contig (THC213329) by The Institute for Genome Research (TIGR). The TIGR annotation indicates that they failed to find any hits in the database for the contig. Note that the TIGR contig is the reverse complement of the LifeSeq contig that we assembled. BLASTN search of Hu-ASP1 against the rat and mouse EST sequences in ZooSeq revealed one homologous EST in each database (Incyte clone 700311523 and IMAGE clone 313341, GenBank accession number W10530, respectively).

TBLASTN searches with the assembled DNA sequence for Hu-ASP1 against both LifeSeqFL and the public EST databases identified a second, related human sequence (Hu-Asp2) represented by a single EST (2696295). Translation of this partial cDNA sequence reveals a single DTG motif which has homology to the active site motif of a bovine aspartyl protease, NM1.

BLAST searches, contig assemblies and multiple sequence alignments were performed using the bioinformatics tools provided with the LifeSeq, LifeSeqFL and LifeSeq Assembled databases from Incyte. Predicted protein motifs were identified using either the ProSite dictionary (Motifs in GCG 9) or the Pfam database.

*Full-length cDNA cloning of Hu-Asp1*

The open reading frame of *C. elegans* gene T18H9.2CE was used to query Incyte LifeSeq and LifeSeq-FL databases and a single electronic assembly referred to as 1863920CE1 was detected. The 5' most cDNA clone in this contig, 1863920, was obtained from Incyte and completely sequenced on both strands. Translation of the open reading frame contained within clone 1863920 revealed the presence of the duplicated aspartyl protease active site motif (DTG/DSG) but the 5' end was incomplete. The remainder of the Hu-Asp1 coding sequence was determined by 5' Marathon RACE analysis using a human placenta Marathon ready cDNA template (Clontech). A 3'-antisense oligonucleotide primer specific for the 5' end of clone 1863920 was paired with the 5'-sense primer specific for the Marathon ready cDNA synthetic adaptor in the PCR. Specific PCR products were directly sequenced by cycle sequencing and the resulting sequence assembled with the sequence of clone 1863920 to yield the complete coding sequence of Hu-Asp-1 (SEQ ID No. 1).

Several interesting features are present in the primary amino acid sequence of Hu-Asp1 (Figure 1, SEQ ID No. 2). The sequence contains a signal peptide (residues 1-20 in SEQ ID No. 2), a pro-segment, and a catalytic domain containing two copies of the aspartyl protease active site motif (DTG/DSG). The spacing between the first and second active site motifs is about 200 residues which should correspond to the expected size of a single, eukaryotic aspartyl protease domain. More interestingly, the sequence contains a predicted transmembrane domain (residues 469-492 in SEQ ID No.2) near its C-terminus which suggests that the protease is anchored in the membrane. This feature is not found in any other aspartyl protease.

#### *Cloning of a full-length Hu-Asp-2 cDNAs:*

As is described above in Example 1, genome wide scan of the *Caenorhabditis elegans* database WormPep12 for putative aspartyl proteases and subsequent mining of human EST databases revealed a human ortholog to the *C. elegans* gene T18H9.2 referred to as Hu-Asp1. The assembled contig for Hu-Asp1 was used to query for human paralogs using the BLAST search tool in human EST databases and a single significant match (2696295CE1) with approximately 60% shared identity was found

in the LifeSeq FL database. Similar queries of either gb105PubEST or the family of human databases available from TIGR did not identify similar EST clones. cDNA clone 2696295, identified by single pass sequence analysis from a human uterus cDNA library, was obtained from Incyte and completely sequenced on both strands. This clone contained an incomplete 1266 bp open-reading frame that encoded a 422 amino acid polypeptide but lacked an initiator ATG on the 5' end. Inspection of the predicted sequence revealed the presence of the duplicated aspartyl protease active site motif DTG/DSG, separated by 194 amino acid residues. Subsequent queries of later releases of the LifeSeq EST database identified an additional ESTs, sequenced from a human astrocyte cDNA library (4386993), that appeared to contain additional 5' sequence relative to clone 2696295. Clone 4386993 was obtained from Incyte and completely sequenced on both strands. Comparative analysis of clone 4386993 and clone 2696295 confirmed that clone 4386993 extended the open-reading frame by 31 amino acid residues including two in-frame translation initiation codons. Despite the presence of the two in-frame ATGs, no in-frame stop codon was observed upstream of the ATG indicating that the 4386993 may not be full-length. Furthermore, alignment of the sequences of clones 2696295 and 4386993 revealed a 75 base pair insertion in clone 2696295 relative to clone 4386993 that results in the insertion of 25 additional amino acid residues in 2696295. The remainder of the Hu-Asp2 coding sequence was determined by 5' Marathon RACE analysis using a human hippocampus Marathon ready cDNA template (Clontech). A 3'-antisense oligonucleotide primer specific for the shared 5'-region of clones 2696295 and 4386993 was paired with the 5'-sense primer specific for the Marathon ready cDNA synthetic adaptor in the PCR. Specific PCR products were directly sequenced by cycle sequencing and the resulting sequence assembled with the sequence of clones 2696295 and 4386993 to yield the complete coding sequence of Hu-Asp2(a) (SEQ ID No. 3) and Hu-Asp2(b) (SEQ ID No. 5), respectively.

Several interesting features are present in the primary amino acid sequence of Hu-Asp2(a) (Figure 2 and SEQ ID No. 4) and Hu-Asp-2(b) (Figure 3, SEQ ID No. 6). Both sequences contain a signal peptide (residues 1-21 in SEQ ID No. 4 and SEQ ID

No. 6), a pro-segment, and a catalytic domain containing two copies of the aspartyl protease active site motif (DTG/DSG). The spacing between the first and second active site motifs is variable due to the 25 amino acid residue deletion in Hu-Asp-2(b) and consists of 168-*versus*-194 amino acid residues, for Hu-Asp2(b) and Hu-Asp-2(a), respectively. More interestingly, both sequences contains a predicted transmembrane domain (residues 455-477 in SEQ ID No.4 and 430-452 in SEQ ID No. 6) near their C-termini which indicates that the protease is anchored in the membrane. This feature is not found in any other aspartyl protease except Hu-Asp1.

### **Example 3**

#### **Molecular cloning of mouse Asp2 cDNA and genomic DNA.**

##### *Cloning and characterization of murine Asp2 cDNA.*

The murine ortholog of Hu-Asp2 was cloned using a combination of cDNA library screening, PCR, and genomic cloning. Approximately 500,000 independent clones from a mouse brain cDNA library were screened using a <sup>32</sup>P-labeled coding sequence probe prepared from Hu-Asp2. Replicate positives were subjected to DNA sequence analysis and the longest cDNA contained the entire 3' untranslated region and 47 amino acids in the coding region. PCR amplification of the same mouse brain cDNA library with an antisense oligonucleotide primer specific for the 5'-most cDNA sequence determined above and a sense primer specific for the 5' region of human Asp2 sequence followed by DNA sequence analysis gave an additional 980 bp of the coding sequence. The remainder of the 5' sequence of murine Asp-2 was derived from genomic sequence (see below).

##### *Isolation and sequence analysis of the murine Asp-2 gene.*

A murine EST sequence encoding a portion of the murine Asp2 cDNA was identified in the GenBank EST database using the BLAST search tool and the Hu-Asp2 coding sequence as the query. Clone g3160898 displayed 88% shared identity to the human sequence over 352 bp. Oligonucleotide primer pairs specific for this region of murine Asp2 were then synthesized and used to amplify regions of the murine gene. Murine genomic DNA, derived from strain 129/SvJ, was amplified in the PCR (25 cycles) using various primer sets specific for murine Asp2 and the

products analyzed by agarose gel electrophoresis. The primer set Zoo-1 and Zoo-4 amplified a 750 bp fragment that contained approximately 600 bp of intron sequence based on comparison to the known cDNA sequence. This primer set was then used to screen a murine BAC library by PCR, a single genomic clone was isolated and this cloned was confirmed contain the murine Asp2 gene by DNA sequence analysis. Shotgun DNA sequencing of this Asp2 genomic clone and comparison to the cDNA sequences of both Hu-Asp2 and the partial murine cDNA sequences defined the full-length sequence of murine Asp2 (SEQ ID No. 7). The predicted amino acid sequence of murine Asp2 (SEQ ID No. 8) showed 96.4% shared identity (GCG BestFit algorithm) with 18/501 amino acid residue substitutions compared to the human sequence (Figure 4). The proteolytic processing of murine Asp2(a) is believed to be analogous to the processing described above for human Asp2(a). In addition, a variant lacking amino acid residues 190-214 of SEQ ID NO: 8 is specifically contemplated as a murine Asp2(b) polypeptide. All forms of murine Asp2(b) gene and protein are intended as aspects of the invention.

#### **Example 4**

##### **Tissue Distribution of Expression of Hu-Asp2 Transcripts**

###### ***Materials and Methods:***

The tissue distribution of expression of Hu-Asp-2 was determined using multiple tissue Northern blots obtained from Clontech (Palo Alto, CA). Incyte clone 2696295 in the vector pINCY was digested to completion with *EcoRI/NotI* and the 1.8 kb cDNA insert purified by preparative agarose gel electrophoresis. This fragment was radiolabeled to a specific activity  $> 1 \times 10^9$  dpm/ $\mu$ g by random priming in the presence of [ $\alpha$ - $^{32}$ P-dATP] ( $>3000$  Ci/mmol, Amersham, Arlington Heights, IL) and Klenow fragment of DNA polymerase I. Nylon filters containing denatured, size fractionated poly A<sup>+</sup> RNAs isolated from different human tissues were hybridized with  $2 \times 10^6$  dpm/ml probe in ExpressHyb buffer (Clontech, Palo Alto, CA) for 1 hour at 68 °C and washed as recommended by the manufacture. Hybridization signals were visualized by autoradiography using BioMax XR film (Kodak, Rochester, NY) with intensifying screens at -80 °C.



### ***Results and Discussion:***

Limited information on the tissue distribution of expression of Hu-Asp-2 transcripts was obtained from database analysis due to the relatively small number of ESTs detected using the methods described above (< 5). In an effort to gain further information on the expression of the Hu-Asp2 gene, Northern analysis was employed to determine both the size(s) and abundance of Hu-Asp2 transcripts. PolyA<sup>+</sup> RNAs isolated from a series of peripheral tissues and brain regions were displayed on a solid support following separation under denaturing conditions and Hu-Asp2 transcripts were visualized by high stringency hybridization to radiolabeled insert from clone 2696295. The 2696295 cDNA probe visualized a constellation of transcripts that migrated with apparent sizes of 3.0kb, 4.4 kb and 8.0 kb with the latter two transcript being the most abundant.

Across the tissues surveyed, Hu-Asp2 transcripts were most abundant in pancreas and brain with lower but detectable levels observed in all other tissues examined except thymus and PBLs. Given the relative abundance of Hu-Asp2 transcripts in brain, the regional expression in brain regions was also established. A similar constellation of transcript sizes were detected in all brain regions examined [cerebellum, cerebral cortex, occipital pole, frontal lobe, temporal lobe and putamen] with the highest abundance in the medulla and spinal cord.

### **Example 5**

#### **Northern Blot Detection of HuAsp-1 and HuAsp-2 Transcripts in Human Cell Lines**

A variety of human cell lines were tested for their ability to produce Hu-Asp1 and Asp2 mRNA. Human embryonic kidney (HEK-293) cells, African green monkey (Cos-7) cells, Chinese hamster ovary (CHO) cells, HELA cells, and the neuroblastoma cell line IMR-32 were all obtained from the ATCC. Cells were cultured in DME containing 10% FCS except CHO cells which were maintained in  $\alpha$ -MEM/10% FCS at 37 °C in 5% CO<sub>2</sub> until they were near confluence. Washed monolayers of cells (3 X 10<sup>7</sup>) were lysed on the dishes and poly A<sup>+</sup> RNA extracted using the Qiagen Oligotex Direct mRNA kit. Samples containing 2  $\mu$ g of poly A<sup>+</sup> RNA from each cell

line were fractionated under denaturing conditions (glyoxal-treated), transferred to a solid nylon membrane support by capillary action, and transcripts visualized by hybridization with random-primed labeled ( $^{32}\text{P}$ ) coding sequence probes derived from either Hu-Asp1 or Hu-Asp2. Radioactive signals were detected by exposure to X-ray film and by image analysis with a PhosphorImager.

The Hu-Asp1 cDNA probe visualized a similar constellation of transcripts (2.6 kb and 3.5 kb) that were previously detected in human tissues. The relative abundance determined by quantification of the radioactive signal was Cos-7 > HEK 292 = HELA > IMR32.

The Hu-Asp2 cDNA probe also visualized a similar constellation of transcripts compared to tissue (3.0 kb, 4.4 kb, and 8.0 kb) with the following relative abundance; HEK 293 > Cos 7 > IMR32 > HELA.

#### **Example 6**

##### **Modification of APP to increase A $\beta$ processing for in vitro screening**

Human cell lines that process A $\beta$  peptide from APP provide a means to screen in cellular assays for inhibitors of  $\beta$ - and  $\gamma$ -secretase. Production and release of A $\beta$  peptide into the culture supernatant is monitored by an enzyme-linked immunosorbent assay (EIA). Although expression of APP is widespread and both neural and non-neuronal cell lines process and release A $\beta$  peptide, levels of endogenous APP processing are low and difficult to detect by EIA. A $\beta$  processing can be increased by expressing in transformed cell lines mutations of APP that enhance A $\beta$  processing. We made the serendipitous observation that addition of two lysine residues to the carboxyl terminus of APP695 increases A $\beta$  processing still further. This allowed us to create a transformed cell line that releases A $\beta$  peptide into the culture medium at the remarkable level of 20,000 pg/ml.

##### ***Materials And Methods***

###### ***Materials:***

Human embryonic kidney cell line 293 (HEK293 cells) were obtained internally. The vector pIRES-EGFP was purchased from Clontech. Oligonucleotides for mutation using the polymerase chain reaction (PCR) were purchased from

Genosys. A plasmid containing human APP695 (SEQ ID No. 9 [nucleotide] and SEQ ID No. 10 [amino acid]) was obtained from Northwestern University Medical School. This was subcloned into pSK (Stratagene) at the *Not*I site creating the plasmid pAPP695.

*Mutagenesis protocol:*

The Swedish mutation (K670N, M671L) was introduced into pAPP695 using the Stratagene Quick Change Mutagenesis Kit to create the plasmid pAPP695NL (SEQ ID No. 11 [nucleotide] and SEQ ID No. 12 [amino acid]). To introduce a di-lysine motif at the C-terminus of APP695, the forward primer #276 5' GACTGACCACTCGACCAGGTTC (SEQ ID No. 47) was used with the "patch" primer #274 5' CGAATTAAATTCCAGCACACTGGCTACTTCTTGTTCTGCATCTCAAAGAAC (SEQ ID No. 48) and the flanking primer #275 CGAATTAAATTCCAGCACACTGGCTA (SEQ ID No. 49) to modify the 3' end of the APP695 cDNA (SEQ ID No. 15 [nucleotide] and SEQ ID No. 16 [amino acid]). This also added a BstX1 restriction site that will be compatible with the BstX1 site in the multiple cloning site of pIRES-EGFP. PCR amplification was performed with a Clontech HF Advantage cDNA PCR kit using the polymerase mix and buffers supplied by the manufacturer. For "patch" PCR, the patch primer was used at 1/20th the molar concentration of the flanking primers. PCR amplification products were purified using a QIAquick PCR purification kit (Qiagen). After digestion with restriction enzymes, products were separated on 0.8% agarose gels and then excised DNA fragments were purified using a QIAquick gel extraction kit (Qiagen).

To reassemble a modified APP695-Sw cDNA, the 5' *Not*I-Bgl2 fragment of the APP695-Sw cDNA and the 3' Bgl2-BstX1 APP695 cDNA fragment obtained by PCR were ligated into pIRES-EGFP plasmid DNA opened at the *Not*I and BstX1 sites. Ligations were performed for 5 minutes at room temperature using a Rapid DNA Ligation kit (Boehringer Mannheim) and transformed into Library Efficiency DH5a Competent Cells (GibcoBRL Life Technologies). Bacterial colonies were screened for inserts by PCR amplification using primers #276 and #275. Plasmid

DNA was purified for mammalian cell transfection using a QIAprep Spin Miniprep kit (Qiagen). The construct obtained was designated pMG125.3 (APPSW-KK, SEQ ID No. 17 [nucleotide] and SEQ ID No. 18 [amino acid]).

#### Mammalian Cell Transfection:

HEK293 cells for transfection were grown to 80% confluence in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum. Cotransfections were performed using LipofectAmine (Gibco-BRL) with 3 µg pMG125.3 DNA and 9 µg pcDNA3.1 DNA per  $10 \times 10^6$  cells. Three days posttransfection, cells were passaged into medium containing G418 at a concentration of 400 µg/ml. After three days growth in selective medium, cells were sorted by their fluorescence.

#### Clonal Selection of 125.3 cells by FACS:

Cell samples were analyzed on an EPICS Elite ESP flow cytometer (Coulter, Hialeah, FL) equipped with a 488 nm excitation line supplied by an air-cooled argon laser. EGFP emission was measured through a 525 nm band-pass filter and fluorescence intensity was displayed on a 4-decade log scale after gating on viable cells as determined by forward and right angle light scatter. Single green cells were separated into each well of one 96 well plate containing growth medium without G418. After a four day recovery period, G418 was added to the medium to a final concentration of 400 µg/ml. After selection, 32% of the wells contained expanding clones. Wells with clones were expanded from the 96 well plate to a 24 well plate and then a 6 well plate with the fastest growing colonies chosen for expansion at each passage. The final cell line selected was the fastest growing of the final six passaged.

This clone, designated 125.3, has been maintained in G418 at 400 µg/ml with passage every four days into fresh medium. No loss of Aβ production or EGFP fluorescence has been seen over 23 passages.

#### *Aβ ELA Analysis (Double Antibody Sandwich ELISA for hAβ 1-40/42)*

Cell culture supernatants harvested 48 hours after transfection were analyzed in a standard Aβ ELA as follows. Human Aβ 1-40 or 1-42 was measured using monoclonal antibody (mAb) 6E10 (Senetek, St. Louis, MO) and biotinylated rabbit antiserum 162 or 164 (New York State Institute for Basic Research, Staten Island, NY) in a double

antibody sandwich ELISA. The capture antibody 6E10 is specific to an epitope present on the N-terminal amino acid residues 1-16 of hA $\beta$ . The conjugated detecting antibodies 162 and 164 are specific for hA $\beta$  1-40 and 1-42, respectively. Briefly, a Nunc Maxisorp 96 well immunoplate was coated with 100  $\mu$ l/well of mAb 6E10 (5 $\mu$ g/ml) diluted in 0.1M carbonate-bicarbonate buffer, pH 9.6 and incubated at 4°C overnight. After washing the plate 3x with 0.01M DPBS (Modified Dulbecco's Phosphate Buffered Saline (0.008M sodium phosphate, 0.002M potassium phosphate, 0.14M sodium chloride, 0.01 M potassium chloride, pH 7.4) from Pierce, Rockford, IL) containing 0.05% of Tween-20 (DPBST), the plate was blocked for 60 minutes with 200  $\mu$ l of 10% normal sheep serum (Sigma) in 0.01M DPBS to avoid non-specific binding. Human A $\beta$  1-40 or 1-42 standards 100  $\mu$ l/well (Bachem, Torrance, CA) diluted, from a 1mg/ml stock solution in DMSO, in culture medium was added after washing the plate, as well as 100  $\mu$ l/well of sample, *e.g.*, conditioned medium of transfected cells.

The plate was incubated for 2 hours at room temperature and 4°C overnight. The next day, after washing the plate, 100  $\mu$ l/well biotinylated rabbit antiserum 162 1:400 or 164 1:50 diluted in DPBST + 0.5% BSA was added and incubated at room temperature for 1 hour, 15 minutes. Following washes, 100  $\mu$ l/well neutravidin-horseradish peroxidase (Pierce, Rockford, IL) diluted 1:10,000 in DPBST was applied and incubated for 1 hour at room temperature. After the last washes 100  $\mu$ l/well of o-phenylenediamine dihydrochloride (Sigma Chemicals, St. Louis, MO) in 50mM citric acid/100mM sodium phosphate buffer (Sigma Chemicals, St. Louis, MO), pH 5.0, was added as substrate and the color development was monitored at 450nm in a kinetic microplate reader for 20 minutes using Soft max Pro software. All standards and samples were run in triplicates. The samples with absorbance values falling within the standard curve were extrapolated from the standard curves using Soft max Pro software and expressed in pg/ml culture medium.

#### **Results:**

Addition of two lysine residues to the carboxyl terminus of APP695 greatly increases A $\beta$  processing in HEK293 cells as shown by transient expression (Table 1).

Addition of the di-lysine motif to APP695 increases A $\beta$  processing to that seen with the APP695 containing the Swedish mutation. Combining the di-lysine motif with the Swedish mutation further increases processing by an additional 2.8 fold.

Cotransformation of HEK293 cells with pMG125.3 and pcDNA3.1 allowed dual selection of transformed cells for G418 resistance and high level expression of EGFP. After clonal selection by FACS, the cell line obtained, produces a remarkable 20,000 pg A $\beta$  peptide per ml of culture medium after growth for 36 hours in 24 well plates. Production of A $\beta$  peptide under various growth conditions is summarized in Table 2.

**TABLE 1**

Release of A $\beta$  peptide into the culture medium 48 hours after transient transfection of HEK293 cells with the indicated vectors containing wildtype or modified APP. Values tabulated are mean + SD and P-value for pairwise comparison using Student's t-test assuming unequal variances.

APP Construct	A $\beta$ 1-40 peptide (pg/ml)	Fold Increase	P-value
pIRES-EGFP vector	147 + 28	1.0	
wt APP695 (142.3)	194 + 15	1.3	0.051
wt APP695-KK (124.1)	424 + 34	2.8	3 x 10 <sup>-5</sup>
APP695-Sw (143.3)	457 + 65	3.1	2 x 10 <sup>-3</sup>
APP695-SwKK (125.3)	1308 + 98	8.9	3 x 10 <sup>-4</sup>

**TABLE 2**

Release of A $\beta$  peptide from HEK125.3 cells under various growth conditions.

Type of Culture Plate	Volume of Medium	Duration of Culture	A $\beta$ 1-40 (pg/ml)	A $\beta$ 1-42 (pg/ml)
24 well plate	400 ul	36 hr	28,036	1,439

**Example 7****Antisense oligomer inhibition of Abeta processing in HEK125.3 cells**

The sequences of Hu-Asp1 and Hu-Asp2 were provided to Sequitur, Inc (Natick, MA) for selection of targeted sequences and design of 2nd generation chimeric antisense oligomers using proprietary technology (Sequitur Ver. D Pat pending #3002). Antisense oligomers Lot# S644, S645, S646 and S647 were targeted against Asp1. Antisense oligomers Lot# S648, S649, S650 and S651 were targeted against Asp2. Control antisense oligomers Lot# S652, S653, S655, and S674 were targeted against an irrelevant gene and antisense oligomers Lot #S656, S657, S658, and S659 were targeted against a second irrelevant gene.

For transfection with the antisense oligomers, HEK125.3 cells were grown to about 50% confluence in 6 well plates in Minimal Essential Medium (MEM) supplemented with 10% fetal calf serum. A stock solution of oligofectin G (Sequitur Inc., Natick, MA) at 2 mg/ml was diluted to 50  $\mu$ g/ml in serum free MEM. Separately, the antisense oligomer stock solution at 100  $\mu$ M was diluted to 800 nM in Opti-MEM (GIBCO-BRL, Grand Island, NY). The diluted stocks of oligofectin G and antisense oligomer were then mixed at a ratio of 1:1 and incubated at room temperature. After 15 minutes incubation, the reagent was diluted 10 fold into MEM containing 10% fetal calf serum and 2 ml was added to each well of the 6 well plate

after first removing the old medium. After transfection, cells were grown in the continual presence of the oligofectin G/antisense oligomer. To monitor A $\beta$  peptide release, 400  $\mu$ l of conditioned medium was removed periodically from the culture well and replaced with fresh medium beginning 24 hours after transfection. A $\beta$  peptides in the conditioned medium were assayed via immunoprecipitation and Western blotting.

Data reported are from culture supernatants harvested 48 hours after transfection.

The 16 different antisense oligomers obtained from Sequitur Inc. were transfected separately into HEK125.3 cells to determine their affect on A $\beta$  peptide processing. Only antisense oligomers targeted against Asp2 significantly reduced Abeta processing by HEK125.3 cells. Both A $\beta$  (1-40) and A $\beta$  (1-42) were inhibited by the same degree. In Table 3, percent inhibition is calculated with respect to untransfected cells. Antisense oligomer reagents giving greater than 50% inhibition are marked with an asterisk. Of the reagents tested, 3 or 4 antisense oligomers targeted against Asp1 gave an average 52% inhibition of A $\beta$ (1-40) processing and 47% inhibition of A $\beta$ (1-42) processing. For Asp2, 4 of 4 antisense oligomers gave greater than 50% inhibition with an average inhibition of 62% of A $\beta$ (1-40) processing and 60% for A $\beta$ (1-42) processing.

**TABLE 3**

Inhibition of A $\beta$  peptide release from HEK125.3 cells treated with antisense oligomers.

Gene Targeted	Antisense Oligomer	Abeta (1-40)	Abeta (1-42)
Asp1-1	S644	62%*	56%*
Asp1-2	S645	41%*	38%*
Asp1-3	S646	52%*	46%*
Asp1-4	S647	6%	25%*
Asp2-1	S648	71%*	67%*
Asp2-2	S649	83%*	76%*
Asp2-3	S650	46%*	50%*
Asp2-4	S651	47%*	46%*



Con1-1	S652	13%	18%
Con1-2	S653	35%	30%
Con1-3	S655	9%	18%
Con1-4	S674	29%	18%
Con2-1	S656	12%	18%
Con2-2	S657	16%	19%
Con2-3	S658	8%	35%
Con2-4	S659	3%	18%

Since HEK293 cells derive from kidney, the experiment was extended to human IMR-32 neuroblastoma cells which express all three APP isoforms and which release A $\beta$  peptides into conditioned medium at measurable levels. [See Neill *et al.*, *J. NeuroSci. Res.*, (1994) 39: 482-93; and Asami-Odaka *et al.*, *Biochem.*, (1995) 34:10272-8.] Essentially identical results were obtained in the neuroblastoma cells as the HEK293 cells. As shown in Table 3B, the pair of Asp2 antisense oligomers reduced Asp2 mRNA by roughly one-half, while the pair of reverse control oligomers lacked this effect (Table 3B).

**Table 3B**

Reduction of A $\beta$ 40 and A $\beta$ 42 in human neuroblastoma IMR-32 cells and mouse neuroblastoma Neuro-2A cells treated with Asp2 antisense and control oligomers as indicated. Oligomers were transfected in quadruplicate cultures. Values tabulated are normalized against cultures treated with oligofectin-G<sup>TM</sup> only (mean + SD, \*\* p<0.001 compared to reverse control oligomer).

		IMR-32 cells		Neuro-2A cells	
	Asp2 mRNA	A $\beta$ 40	A $\beta$ 42	A $\beta$ 40	A $\beta$ 42
Asp2-1A	-75%	-49 + 2%**	-42 + 14%**	-70 + 7%**	-67 + 2%**
Asp2-1R	0.16	-0 + 3%	21.26	-9 + 15%	1.05
Asp2-2A	-39%	-43 + 3%**	-44 + 18%**	-61 + 12%**	-61 + 12%**
Asp2-2R	0.47	12.2	19.22	6.15	-8 + 10%

Together with the reduction in Asp2 mRNA there was a concomitant reduction in the release of A $\beta$ 40 and A $\beta$ 42 peptides into the conditioned medium. Thus, Asp2 functions directly or indirectly in a human kidney and a human neuroblastoma cell line to facilitate the processing of APP into A $\beta$  peptides. Molecular cloning of the mouse Asp2 cDNA revealed a high degree of homology to human (>96% amino acid identity, see Example 3), and indeed, complete nucleotide identity at the sites targeted by the Asp2-1A and Asp2-2A antisense oligomers. Similar results were obtained in mouse Neuro-2a cells engineered to express APP-Sw-KK. The Asp2 antisense oligomers reduced release of A $\beta$  peptides into the medium while the reverse control oligomers did not (Table 3B). Thus, the three antisense experiments with HEK293,

IMR-32 and Neuro-2a cells indicate that Asp2 acts directly or indirectly to facilitate A $\beta$  processing in both somatic and neural cell lines.

#### **Example 8**

##### **Demonstration of Hu-Asp2 $\beta$ -Secretase Activity in Cultured Cells**

Several mutations in APP associated with early onset Alzheimer's disease have been shown to alter A $\beta$  peptide processing. These flank the – and C-terminal cleavage sites that release A $\beta$  from APP. These cleavage sites are referred to as the  $\beta$ -secretase and  $\gamma$ -secretase cleavage sites, respectively. Cleavage of APP at the  $\beta$ -secretase site creates a C-terminal fragment of APP containing 99 amino acids of 11,145 daltons molecular weight. The Swedish KM $\rightarrow$ NL mutation immediately upstream of the  $\beta$ -secretase cleavage site causes a general increase in production of both the 1-40 and 1-42 amino acid forms of A $\beta$  peptide. The London VF mutation (V717 $\rightarrow$ F in the APP770 isoform) has little effect on total A $\beta$  peptide production, but appears to preferentially increase the percentage of the longer 1-42 amino acid form of A $\beta$  peptide by affecting the choice of  $\beta$ -secretase cleavage site used during APP processing. Thus, we sought to determine if these mutations altered the amount and type of A $\beta$  peptide produced by cultured cells cotransfected with a construct directing expression of Hu-Asp2.

Two experiments were performed which demonstrate Hu-Asp2  $\beta$ -secretase activity in cultured cells. In the first experiment, treatment of HEK125.3 cells with antisense oligomers directed against Hu-Asp2 transcripts as described in Example 7 was found to decrease the amount of the C-terminal fragment of APP created by  $\beta$ -secretase cleavage (CTF99) (Figure 9). This shows that Hu-Asp2 acts directly or indirectly to facilitate  $\beta$ -secretase cleavage. In the second experiment, increased expression of Hu-Asp2 in transfected mouse Neuro2A cells is shown to increase accumulation of the CTF99  $\beta$ -secretase cleavage fragment (Figure 10). This increase is seen most easily when a mutant APP-KK clone containing a C-terminal di-lysine motif is used for transfection. A further increase is seen when Hu-Asp2 is cotransfected with APP-Sw-KK containing the Swedish mutation KM  $\rightarrow$  NL. The Swedish mutation is known to increase cleavage of APP by the  $\beta$ -secretase.

A second set of experiments demonstrate Hu-Asp2 facilitates  $\gamma$ -secretase activity in cotransfection experiments with human embryonic kidney HEK293 cells. Cotransfection of Hu-Asp2 with an APP-KK clone greatly increases production and release of soluble A $\beta$ 1-40 and A $\beta$ 1-42 peptides from HEK293 cells. There is a proportionately greater increase in the release of A $\beta$ 1-42. A further increase in production of A $\beta$ 1-42 is seen when Hu-Asp2 is cotransfected with APP-VF (SEQ ID No. 13 [nucleotide] and SEQ ID No. 14 [amino acid]) or APP-VF-KK (SEQ ID No. 19 [nucleotide] and SEQ ID No. 20 [amino acid]) clones containing the London mutation V717→F. The V717→F mutation is known to alter cleavage specificity of the APP  $\gamma$ -secretase such that the preference for cleavage at the A $\beta$ 42 site is increased. Thus, Asp2 acts directly or indirectly to facilitate  $\gamma$ -secretase processing of APP at the  $\beta$ 42 cleavage site.

#### *Materials*

Antibodies 6E10 and 4G8 were purchased from Senetek (St. Louis, MO). Antibody 369 was obtained from the laboratory of Paul Greengard at the Rockefeller University. Antibody C8 was obtained from the laboratory of Dennis Selkoe at the Harvard Medical School and Brigham and Women's Hospital.

#### *APP Constructs used*

The APP constructs used for transfection experiments comprised the following APP: wild-type APP695 (SEQ ID No. 9 and No. 10)

APP-Sw: APP695 containing the Swedish KM→NL mutation (SEQ ID No. 11 and No. 12), wherein the lysine (K) at residue 595 of APP695 is changed to asparagine (N) and the methionine (M) at residue 596 of APP695 is changed to leucine (L).),

APP-VF: APP695 containing the London V→F mutation (SEQ ID Nos. 13 & 14) (Affected residue 717 of the APP770 isoform corresponds with residue 642 of the APP695 isoform. Thus, APP-VF as set in SEQ ID NO: 14 comprises the APP695 sequence, wherein the valine (V) at residue 642 is changed to phenylalanine (F).)

APP-KK: APP695 containing a C-terminal KK motif (SEQ ID Nos. 15 & 16),

APP-Sw-KK: APP695-Sw containing a C-terminal KK motif (SEQ ID No. 17 & 18),

APP-VF-KK: APP695-VF containing a C-terminal KK motif (SEQ ID Nos. 19 & 20).

These were inserted into the vector pIRES-EGFP (Clontech, Palo Alto CA) between the *Not*I and *Bst*XI sites using appropriate linker sequences introduced by PCR.

*Transfection of antisense oligomers or plasmid DNA constructs in HEK293 cells, HEK125.3 cells and Neuro-2A cells,*

Human embryonic kidney HEK293 cells and mouse Neuro-2a cells were transfected with expression constructs using the Lipofectamine Plus reagent from Gibco/BRL. Cells were seeded in 24 well tissue culture plates to a density of 70-80% confluence. Four wells per plate were transfected with 2 µg DNA (3:1, APP:cotransfectant), 8 µl Plus reagent, and 4 µl Lipofectamine in OptiMEM. OptiMEM was added to a total volume of 1 ml, distributed 200 µl per well and incubated 3 hours. Care was taken to hold constant the ratios of the two plasmids used for cotransfection as well as the total amount of DNA used in the transfection. The transfection media was replaced with DMEM, 10%FBS, NaPyruvate, with antibiotic/antimycotic and the cells were incubated under normal conditions (37°C, 5% CO<sub>2</sub>) for 48 hours. The conditioned media were removed to polypropylene tubes and stored at -80°C until assayed for the content of Aβ1-40 and Aβ1-42 by EIA as described in the preceding examples. Transfection of antisense oligomers into HEK125.3 cells was as described in Example 7.

*Preparation of cell extracts, Western blot protocol*

Cells were harvested after being transfected with plasmid DNA for about 60 hours. First, cells were transferred to 15-ml conical tube from the plate and centrifuged at 1,500 rpm for 5 minutes to remove the medium. The cell pellets were washed once with PBS. We then lysed the cells with lysis buffer (10 mM HEPES, pH 7.9, 150 mM NaCl, 10% glycerol, 1 mM EGTA, 1 mM EDTA, 0.1 mM sodium vanadate and 1% NP-40). The lysed cell mixtures were centrifuged at 5000 rpm and

the supernatant was stored at -20°C as the cell extracts. Equal amounts of extracts from HEK125.3 cells transfected with the Asp2 antisense oligomers and controls were precipitated with antibody 369 that recognizes the C-terminus of APP and then CTF99 was detected in the immunoprecipitate with antibody 6E10. The experiment was repeated using C8, a second precipitating antibody that also recognizes the C-terminus of APP. For Western blot of extracts from mouse Neuro-2a cells cotransfected with Hu-Asp2 and APP-KK, APP-Sw-KK, APP-VF-KK or APP-VF, equal amounts of cell extracts were electrophoresed through 4-10% or 10-20% Tricine gradient gels (NOVEX, San Diego, CA). Full length APP and the CTF99  $\beta$ -secretase product were detected with antibody 6E10.

### ***Results***

Transfection of HEK125.3 cells with Asp2-1 or Asp2-2 antisense oligomers reduces production of the CTF  $\beta$ -secretase product in comparison to cells similarly transfected with control oligomers having the reverse sequence (Asp2-1 reverse & Asp2-2 reverse), see Figure 9. Correspondingly, cotransfection of Hu-Asp2 into mouse Neuro-2a cells with the APP-KK construct increased the formation of CTF99. (See Fig. 10.) This was further increased if Hu-Asp2 was coexpressed with APP-Sw-KK, a mutant form of APP containing the Swedish KM→NL mutation that increases  $\beta$ -secretase processing.

Effects of Asp2 on the production of Ab peptides from endogenously expressed APP isoforms were assessed in HEK293 cells transfected with a construct expressing Asp2 or with the empty vector after selection of transformants with the antibiotic G418. A $\beta$ 40 production was increased in cells transformed with the Asp2 construct in comparison to those transformed with the empty vector DNA. A $\beta$ 40 levels in conditioned medium collected from the Asp2 transformed and control cultures was  $424 \pm 45$  pg/ml and  $113 \pm 58$  pg/ml, respectively ( $p < 0.001$ ). A $\beta$ 42 release was below the limit of detection by the EIA, while the release of sAPP $\alpha$  was unaffected,  $112 \pm 8$  ng/ml versus  $111 \pm 40$  ng/ml. This further indicates that Asp2 acts directly or indirectly to facilitate the processing and release of A $\beta$  from endogenously expressed APP.

Co-transfection of Hu-Asp2 with APP has little effect on A $\beta$ 40 production but increases A $\beta$ 42 production above background (Table 4). Addition of the di-lysine motif to the C-terminus of APP increases A $\beta$  peptide processing about two fold, although A $\beta$ 40 and A $\beta$ 42 production remain quite low (352 pg/ml and 21 pg/ml, respectively). Cotransfection of Asp2 with APP-KK further increases both A $\beta$ 40 and A $\beta$ 42 production.

The APP V717→F mutation has been shown to increase  $\gamma$ -secretase processing at the A $\beta$ 42 cleavage site. Cotransfection of Hu-Asp2 with the APP-VF or APP-VF-KK constructs increased A $\beta$ 42 production (a two fold increase with APP-VF and a four-fold increase with APP-VF-KK, Table 4), but had mixed effects on A $\beta$ 40 production (a slight decrease with APP-VF, and a two fold increase with APP-VF-KK in comparison to the pcDNA cotransfection control. Thus, the effect of Asp2 on A $\beta$ 42 production was proportionately greater leading to an increase in the ratio of A $\beta$ 42/total Ab. Indeed, the ratio of A $\beta$ 42/total A $\beta$  reaches a very high value of 42% in HEK293 cells cotransfected with Hu-Asp2 and APP-VF-KK.

**Table 4**

Results of cotransfecting Hu-Asp2 or pcDNA plasmid DNA with various APP constructs containing the V717→F mutation that modifies  $\gamma$ -secretase processing. Cotransfection with Asp2 consistently increases the ratio of A $\beta$ 42/total A $\beta$ . Values tabulated are A $\beta$  peptide pg/ml.

pcDNA Cotransfection			Asp2 Cotransfection		
A $\beta$ 40	A $\beta$ 42	A $\beta$ 42/Tot al	A $\beta$ 40	A $\beta$ 42	A $\beta$ 42/Tot al
<hr/>					
<hr/>					
<hr/>					

APP	192 $\pm$ 1	<4	<2%	188 $\pm$ 40	8 $\pm$ 10	3.9%
	8					
APP-VF	118 $\pm$ 1	15 $\pm$ 19	11.5%	85 $\pm$ 7	24 $\pm$ 12	22.4%
	5					
APP-KK	352 $\pm$ 2	21 $\pm$ 6	5.5%	1062 $\pm$ 101	226 $\pm$ 4	17.5%
	4				9	
APP-VF-K	230 $\pm$ 3	88 $\pm$ 24	27.7%	491 $\pm$ 35	355 $\pm$ 3	42%
K	1				6	

### Example 9

#### Bacterial expression of human Asp2(a)

##### *Expression of recombinant Hu-Asp2(a) in E. coli.*

Hu-Asp2(a) can be expressed in *E. coli* after addition of N-terminal sequences such as a T7 tag (SEQ ID No. 21 and No. 22) or a T7 tag followed by a caspase 8 leader sequence (SEQ ID No. 23 and No. 24). Alternatively, reduction of the GC content of the 5' sequence by site directed mutagenesis can be used to increase the yield of Hu-Asp2 (SEQ ID No. 25 and No. 26). In addition, Asp2(a) can be engineered with a proteolytic cleavage site (SEQ ID No. 27 and No. 28). To produce a soluble protein after expression and refolding, deletion of the transmembrane domain and cytoplasmic tail, or deletion of the membrane proximal region, transmembrane domain, and cytoplasmic tail is preferred. Any materials (vectors, host cells, etc.) and methods described herein to express Hu-Asp2(a) should in principle be equally effective for expression of Hu-Asp2(b).



## Methods

PCR with primers containing appropriate linker sequences was used to assemble fusions of Asp2(a) coding sequence with N-terminal sequence modifications including a T7 tag (SEQ ID Nos. 21 and 22) or a T7-caspase 8 leader (SEQ ID Nos. 23 and 24). These constructs were cloned into the expression vector pet23a(+) [Novagen] in which a T7 promoter directs expression of a T7 tag preceding a sequence of multiple cloning sites. To clone Hu-Asp2 sequences behind the T7 leader of pet23a+, the following oligonucleotides were used for amplification of the selected Hu-Asp2(a) sequence: #553=GTGGATCCACCCAGCACGGCATCCGGCTG (SEQ ID No. 35), #554=GAAAGCTTTCATGACTCATCTGTCTGTGGAATGTTG (SEQ ID No. 36) which placed BamHI and HindIII sites flanking the 5' and 3' ends of the insert, respectively. The Asp2(a) sequence was amplified from the full length Asp2(a) cDNA cloned into pcDNA3.1 using the Advantage-GC cDNA PCR [Clontech] following the manufacturer's supplied protocol using annealing & extension at 68°C in a two-step PCR cycle for 25 cycles. The insert and vector were cut with BamHI and HindIII, purified by electrophoresis through an agarose gel, then ligated using the Rapid DNA Ligation kit [Boehringer Mannheim]. The ligation reaction was used to transform the *E. coli* strain JM109 (Promega) and colonies were picked for the purification of plasmid (Qiagen, Qiaprep minispin) and DNA sequence analysis. For inducible expression using induction with isopropyl b-D-thiogalactopyranoside (IPTG), the expression vector was transferred into *E. coli* strain BL21 (Statagene). Bacterial cultures were grown in LB broth in the presence of ampicillin at 100 ug/ml, and induced in log phase growth at an OD600 of 0.6-1.0 with 1 mM IPTG for 4 hour at 37°C. The cell pellet was harvested by centrifugation.

To clone Hu-Asp2 sequences behind the T7 tag and caspase leader (SEQ ID Nos. 23 and 24), the construct created above containing the T7-Hu-Asp2 sequence (SEQ ID Nos. 21 and 22) was opened at the BamHI site, and then the phosphorylated caspase 8 leader oligonucleotides #559=GATCGATGACTATCTCTGACTCTCCGCGTGAACAGGACG (SEQ ID No. 37), #560=GATCCGTCCTGTTCACGCGGAGAGTCAGAGATAGTCATC (SEQ

ID No. 38) were annealed and ligated to the vector DNA. The 5' overhang for each set of oligonucleotides was designed such that it allowed ligation into the BamHI site but not subsequent digestion with BamHI. The ligation reaction was transformed into JM109 as above for analysis of protein expression after transfer to *E. coli* strain BL21.

In order to reduce the GC content of the 5' terminus of asp2(a), a pair of antiparallel oligos were designed to change degenerate codon bases in 15 amino acid positions from G/C to A/T (SEQ ID Nos. 25 and 26). The new nucleotide sequence at the 5' end of asp2 did not change the encoded amino acid and was chosen to optimize *E. Coli* expression. The sequence of the sense linker is 5'

CGGCATCCGGCTGCCCCCTGCGTAGCGGTCTGGGTGGTGCTCCACTGGGTCT  
GCGTCTGCCCCGGGAGACCGACGAA G 3' (SEQ ID No. 39). The sequence of the antisense linker is : 5'

CTTCGTCTGGTCTCCCGGGGCAGACGCAGACCCAGTGGAGCACCACCCAGA  
CCGCTACGCAGGGGGCAGCCGGATGCCG 3' (SEQ ID No. 40). After annealing the phosphorylated linkers together in 0.1 M NaCl-10 mM Tris, pH 7.4 they were ligated into unique Cla I and Sma I sites in Hu-Asp2 in the vector pTAC. For inducible expression using induction with isopropyl b-D-thiogalactopyranoside (IPTG), bacterial cultures were grown in LB broth in the presence of ampicillin at 100 ug/ml, and induced in log phase growth at an OD600 of 0.6-1.0 with 1 mM IPTG for 4 hour at 37°C. The cell pellet was harvested by centrifugation.

To create a vector in which the leader sequences can be removed by limited proteolysis with caspase 8 such that this liberates a Hu-Asp2 polypeptide beginning with the N-terminal sequence GSFV (SEQ ID Nos. 27 and 28), the following procedure was followed. Two phosphorylated oligonucleotides containing the caspase 8 cleavage site IETD, #571=5'

GATCGATGACTATCTCTGACTCTCCGCTGGACTCTGGTATCGAAACCGACG  
(SEQ ID No. 41) and #572=

GATCCGTCGGTTTCGATACCAGAGTCCAGCGGAGAGTCAGAGATAGTCAT  
C (SEQ ID No. 42) were annealed and ligated into pET23a+ that had been opened



193mMhos with water. After the pellet was dispersed, an additional amount of the KCl solution was added, bringing the total volume to 500 ml. This suspension was homogenized further for about 3 minutes at 5000 rpm using the same probe. The mixture was then passed through a Rannie high-pressure homogenizer at 10,000psi.

In all cases, the pellet material was carried forward, while the soluble fraction was discarded. The resultant solution was centrifuged in a GSA rotor for 1 hour at 12,500 rpm. The pellet was resuspended in the same solution (without the DTT) using the same tissue homogenizer probe at 2,000 rpm. After homogenizing for 5 minutes at 3000 rpm, the volume was adjusted to 500ml with the same solution, and spun for 1 hour at 12,500 rpm. The pellet was then resuspended as before, but this time the final volume was adjusted to 1.5L with the same solution prior to homogenizing for 5 minutes. After centrifuging at the same speed for 30 minutes, this procedure was repeated. The pellet was then resuspended into about 150ml of cold water, pooling the pellets from the six centrifuge tubes used in the GSA rotor. The pellet was homogenized for 5 minutes at 3,000 rpm, volume adjusted to 250ml with cold water, then spun for 30 minutes. Weight of the resultant pellet was 17.75g.

Summary: Lysis of bacterial pellet in KCl solution, followed by centrifugation in a GSA rotor was used to initially prepare the pellet. The same solution was then used an additional three times for resuspension/homogenization. A final water wash/homogenization was then performed to remove excess KCl and EDTA.

*Solubilization of Recombinant Hu-Asp2(a):*

A ratio of 9-10ml/gram of pellet was utilized for solubilizing the rHuAsp2L from the pellet previously described. 17.75g of pellet was thawed, and 150ml of 8M guanidine HCl, 5mM  $\beta$ ME, 0.1% DEA, was added. 3M Tris was used to titrate the pH to 8.6. The pellet was initially resuspended into the guanidine solution using a 20 mm tissue homogenizer probe at 1000 rpm. The mixture was then stirred at 4°C for 1 hour prior to centrifugation at 12,500 rpm for 1 hour in GSA rotor. The resultant supernatant was then centrifuged for 30 minutes at 40,000 x g in an SS-34 rotor. The final supernatant was then stored at -20°C, except for 50 ml.

*Immobilized Nickel Affinity Chromatography of Solubilized Recombinant Hu-Asp2(a):*

The following solutions were utilized:

A) 6M Guanidine HCl, 0.1M NaP, pH 8.0, 0.01M Tris, 5mM  $\beta$ ME, 0.5mM

Imidazole

A') 6M Urea, 20mM NaP, pH 6.80, 50mM NaCl

B') 6M Urea, 20mM NaP, pH 6.20, 50mM NaCl, 12mM Imidazole

C') 6M Urea, 20mM NaP, pH 6.80, 50mM NaCl, 300mM Imidazole

Note: Buffers A' and C' were mixed at the appropriate ratios to give intermediate concentrations of Imidazole.

The 50ml of solubilized material was combined with 50ml of buffer A prior to adding to 100-125ml Qiagen Ni-NTA SuperFlow (pre-equilibrated with buffer A) in a 5 x 10cm Bio-Rad econo column. This was shaken gently overnight at 4°C in the cold room.

*Chromatography Steps:*

Drained the resultant flow through.

Washed with 50ml buffer A (collecting into flow through fraction)

Washed with 250ml buffer A (wash 1)

Washed with 250ml buffer A (wash 2)

Washed with 250ml buffer A'

Washed with 250ml buffer B'

Washed with 250ml buffer A'

Eluted with 250ml 75mM Imidazole

Eluted with 250ml 150mM Imidazole (150-1)

Eluted with 250ml 150mM Imidazole (150-2)

Eluted with 250ml 300mM Imidazole (300-1)

Eluted with 250ml 300mM Imidazole (300-2)

Eluted with 250ml 300mM Imidazole (300-3)

### *Chromatography Results:*

The Hu-Asp(a) eluted at 75mM Imidazole through 300mM Imidazole. The 75mM fraction, as well as the first 150mM Imidazole (150-1) fraction contained contaminating proteins as visualized on Coomassie Blue stained gels. Therefore, fractions 150-2 and 300-1 will be utilized for refolding experiments since they contained the greatest amount of protein as visualized on a Coomassie Blue stained gel.

### *Refolding Experiments of Recombinant Hu-Asp2(a):*

#### *Experiment 1:*

Forty ml of 150-2 was spiked with 1M DTT, 3M Tris, pH 7.4 and DEA to a final concentration of 6mM, 50mM, and 0.1% respectively. This was diluted suddenly (while stirring) with 200ml of (4°C) cold 20mM NaP, pH 6.8, 150mM NaCl. This dilution gave a final Urea concentration of 1M. This solution remained clear, even if allowed to set open to the air at room temperature (RT) or at 4°C .

After setting open to the air for 4-5 hours at 4°C, this solution was then dialyzed overnight against 20 mM NaP, pH 7.4, 150 mM NaCl, 20% glycerol. This method effectively removes the urea in the solution without precipitation of the protein.

#### *Experiment 2:*

Some of the 150-2 eluate was concentrated 2x on an Amicon Centriprep, 10,000 MWCO, then treated as in Experiment 1. This material also stayed in solution, with no visible precipitation.

#### *Experiment 3:*

89ml of the 150-2 eluate was spiked with 1M DTT, 3M Tris, pH 7.4 and DEA to a final concentration of 6mM, 50mM, and 0.1% respectively. This was diluted suddenly (while stirring) with 445 ml of (4°C) cold 20 mM NaP, pH 6.8, 150 mM NaCl. This solution appeared clear, with no apparent precipitation. The solution was removed to RT and stirred for 10 minutes prior to adding MEA to a final

concentration of 0.1 mM. This was stirred slowly at RT for 1 hour. Cystamine and CuSO<sub>4</sub> were then added to final concentrations of 1 mM and 10 μM respectively. The solution was stirred slowly at RT for 10 minutes prior to being moved to the 4°C cold room and shaken slowly overnight, open to the air.

The following day, the solution (still clear, with no apparent precipitation) was centrifuged at 100,000 x g for 1 hour. Supernatants from multiple runs were pooled, and the bulk of the stabilized protein was dialyzed against 20mM NaP, pH 7.4, 150 mM NaCl, 20% glycerol. After dialysis, the material was stored at -20°C.

Some (about 10 ml) of the protein solution (still in 1M Urea) was saved back for biochemical analyses, and frozen at -20°C for storage.

### **Example 10**

#### **Expression of Hu-Asp2 and Derivatives in Insect Cells**

Any materials (vectors, host cells, etc.) and methods that are useful to express Hu-Asp2(a) should in principle be equally effective for expression of Hu-Asp2(b).

#### ***Expression by baculovirus infection.***

The coding sequence of Hu-Asp2(a) and Hu-ASp2(b) and several derivatives were engineered for expression in insect cells using the PCR. For the full-length sequence, a 5'-sense oligonucleotide primer that modified the translation initiation site to fit the Kozak consensus sequence was paired with a 3'-antisense primer that contains the natural translation termination codon in the Hu-Asp2 sequence. PCR amplification of the pcDNA3.1(hygro)/Hu-Asp2(a) template was used to prepare two derivatives of Hu-Asp2(a) or Hu-Asp(b) that delete the C-terminal transmembrane domain (SEQ ID Nos. 29-30 and 50-51, respectively) or delete the transmembrane domain and introduce a hexa-histidine tag at the C-terminus (SEQ ID Nos. 31-32 and 52-53) respectively, were also engineered using PCR. The same 5'-sense oligonucleotide primer described above was paired with either a 3'-antisense primer that (1) introduced a translation termination codon after codon 453 (SEQ ID No. 3) or (2) incorporated a hexa-histidine tag followed by a translation termination codon in the PCR using pcDNA3.1(hygro)/Hu-Asp-2(a) as the template. In all cases, the PCR reactions were performed amplified for 15 cycles using *Pwo*I DNA polymerase

(Boehringer-Mannheim) as outlined by the supplier. The reaction products were digested to completion with *Bam*HI and *Not*I and ligated to *Bam*HI and *Not*I digested baculovirus transfer vector pVL1393 (Invitrogen). A portion of the ligations was used to transform competent *E. coli* DH5 $\alpha$  cells followed by antibiotic selection on LB-Amp. Plasmid DNA was prepared by standard alkaline lysis and banding in CsCl to yield the baculovirus transfer vectors pVL1393/Asp2(a), pVL1393/Asp2(a) $\Delta$ TM and pVL1393/Asp2(a) $\Delta$ TM(His) $_6$ . Creation of recombinant baculoviruses and infection of sf9 insect cells was performed using standard methods.

#### *Expression by transfection*

Transient and stable expression of Hu-Asp2(a) $\Delta$ TM and Hu-Asp2(a) $\Delta$ TM(His) $_6$  in High 5 insect cells was performed using the insect expression vector pIZ/V5-His. The DNA inserts from the expression plasmids vectors pVL1393/Asp2(a), pVL1393/Asp2(a) $\Delta$ TM and pVL1393/Asp2(a) $\Delta$ TM(His) $_6$  were excised by double digestion with *Bam*HI and *Not*I and subcloned into *Bam*HI and *Not*I digested pIZ/V5-His using standard methods. The resulting expression plasmids, referred to as pIZ/Hu-Asp2 $\Delta$ TM and pIZ/Hu-Asp2 $\Delta$ TM(His) $_6$ , were prepared as described above.

For transfection, High 5 insect cells were cultured in High Five serum free medium supplemented with 10  $\mu$ g/ml gentamycin at 27 °C in sealed flasks. Transfections were performed using High five cells, High five serum free media supplemented with 10  $\mu$ g/ml gentamycin, and InsectinPlus liposomes (Invitrogen, Carlsbad, CA) using standard methods.

For large scale transient transfections,  $1.2 \times 10^7$  high five cells were plated in a 150 mm tissue culture dish and allowed to attach at room temperature for 15-30 minutes. During the attachment time the DNA/ liposome mixture was prepared by mixing 6 ml of serum free media, 60  $\mu$ g Hu-Asp2(a) $\Delta$ TM/pIZ (+/- His) DNA and 120  $\mu$ l of Insectin Plus and incubating at room temperature for 15 minutes. The plating media was removed from the dish of cells and replaced with the DNA/liposome mixture for 4 hours at room temperature with constant rocking at 2 rpm. An additional 6 ml of media was added to the dish prior to incubation for 4 days at 27 °C



in a humid incubator. Four days post transfection the media was harvested, clarified by centrifugation at 500 x g, assayed for Hu-Asp2(a) expression by Western blotting. For stable expression, the cells were treated with 50 µg/ml Zeocin and the surviving pool used to prepared clonal cells by limiting dilution followed by analysis of the expression level as noted above.

*Purification of Hu-Asp2(a)ΔTM and Hu-Asp2(a)ΔTM(His)<sub>6</sub>*

Removal of the transmembrane segment from Hu-Asp2(a) resulted in the secretion of the polypeptide into the culture medium. Following protein production by either baculovirus infection or transfection, the conditioned medium was harvested, clarified by centrifugation, and dialyzed against Tris-HCl (pH 8.0). This material was then purified by successive chromatography by anion exchange (Tris-HCl, pH 8.0) followed by cation exchange chromatography (Acetate buffer at pH 4.5) using NaCl gradients. The elution profile was monitored by (1) Western blot analysis and (2) by activity assay using the peptide substrate described in Example 12. For the Hu-Asp2(a)ΔTM(His)<sub>6</sub>, the conditioned medium was dialyzed against Tris buffer (pH 8.0) and purified by sequential chromatography on IMAC resin followed by anion exchange chromatography.

Amino-terminal sequence analysis of the purified Hu-Asp2(a)ΔTM(His)<sub>6</sub> protein revealed that the signal peptide had been cleaved [TQHGIRLPLR, corresponding to SEQ ID NO: 32, residues 22-3].

**Example 11**

**Expression of Hu-Asp2(a) and Hu-Asp(b) in CHO cells**

The materials (vectors, host cells, etc.) and methods described herein for expression of Hu-Asp2(a) are intended to be equally applicable for expression of Hu-Asp2(b).

*Heterologous expression of Hu-Asp-2(a) in CHO-K1 cells*

The entire coding sequence of Hu-Asp2(a) was cloned into the mammalian expression vector pcDNA3.1(+)/Hygro (Invitrogen, Carlsbad, CA) which contains the CMV immediate early promoter and bGH polyadenylation signal to drive over expression. The expression plasmid, pcDNA3.1(+)/Hygro/Hu-Asp2(a), was prepared

by alkaline lysis and banding in CsCl and completely sequenced on both strands to verify the integrity of the coding sequence.

Wild-type Chinese hamster ovary cells (CHO-K1) were obtained from the ATCC. The cells were maintained in monolayer cultures in  $\alpha$ -MEM containing 10% FCS at 37°C in 5% CO<sub>2</sub>. Two 100 mm dishes of CHO-K1 cells (60% confluent) were transfected with pcDNA3.1(+)/Hygro alone (mock) or pcDNA3.1(+)/Hygro/Hu-Asp2(a) or pcDNA3.1(+)/Hygro/Hu-Asp2(b) using the cationic liposome DOTAP as recommended by the supplier (Roche, Indianapolis, IN). The cells were treated with the plasmid DNA/liposome mixtures for 15 hours and then the medium replaced with growth medium containing 500 Units/ml hygromycin B. In the case of pcDNA3.1(+)/Hygro/Hu-Asp2(a) or (b) transfected CHO-K1 cells, individual hygromycin B-resistant cells were cloned by limiting dilution. Following clonal expansion of the individual cell lines, expression of Hu-Asp2(a) or Hu-Asp2(b) protein was assessed by Western blot analysis using a polyclonal rabbit antiserum raised against recombinant Hu-Asp2 prepared by expression in *E. coli*. Near confluent dishes of each cell line were harvested by scraping into PBS and the cells recovered by centrifugation. The cell pellets were resuspended in cold lysis buffer (25 mM Tris-HCl (pH 8.0)/5 mM EDTA) containing protease inhibitors and the cells lysed by sonication. The soluble and membrane fractions were separated by centrifugation (105,000 x g, 60 min) and normalized amounts of protein from each fraction were then separated by SDS-PAGE. Following electrotransfer of the separated polypeptides to PVDF membranes, Hu-Asp-2(a) or Hu-Asp2(b) protein was detected using rabbit anti-Hu-Asp2 antiserum (1/1000 dilution) and the antibody-antigen complexes were visualized using alkaline phosphatase conjugated goat anti-rabbit antibodies (1/2500). A specific immunoreactive protein with an apparent Mr value of 65 kDa was detected in pcDNA3.1(+)/Hygro/Hu-Asp2 transfected cells and not mock-transfected cells. Also, the Hu-Asp2 polypeptide was only detected in the membrane fraction, consistent with the presence of a signal peptide and single transmembrane domain in the predicted sequence. Based on this

analysis, clone #5 had the highest expression level of Hu-Asp2(a) protein and this production cell lines was scaled up to provide material for purification.

*Purification of recombinant Hu-Asp-2(a) from CHO-K1/Hu-Asp2 clone #5*

In a typical purification, clone #5 cell pellets derived from 20 150 mm dishes of confluent cells, were used as the starting material. The cell pellets were resuspended in 50 ml cold lysis buffer as described above. The cells were lysed by polytron homogenization (2 x 20 sec) and the lysate centrifuged at 338,000 x g for 20 minutes. The membrane pellet was then resuspended in 20 ml of cold lysis buffer containing 50 mM  $\beta$ -octylglucoside followed by rocking at 4 °C for 1 hour. The detergent extract was clarified by centrifugation at 338,000 x g for 20 minutes and the supernatant taken for further analysis.

The  $\beta$ -octylglucoside extract was applied to a Mono Q anion exchange column that was previously equilibrated with 25 mM Tris-HCl (pH 8.0)/50 mM  $\beta$ -octylglucoside. Following sample application, the column was eluted with a linear gradient of increasing NaCl concentration (0-1.0 M over 30 minutes) and individual fractions assayed by Western blot analysis and for  $\beta$ -secretase activity (see below). Fractions containing both Hu-Asp-2(a) immunoreactivity and  $\beta$ -secretase activity were pooled and dialyzed against 25 mM NaOAc (pH 4.5)/50 mM  $\beta$ -octylglucoside. Following dialysis, precipitated material was removed by centrifugation and the soluble material chromatographed on a MonoS cation exchange column that was previously equilibrated in 25 mM NaOAc (pH 4.5)/ 50 mM  $\beta$ -octylglucoside. The column was eluted using a linear gradient of increasing NaCl concentration (0-1.0 M over 30 minutes) and individual fractions assayed by Western blot analysis and for  $\beta$ -secretase activity. Fractions containing both Hu-Asp2 immunoreactivity and  $\beta$ -secretase activity were combined and determined to be >95% pure by SDS-PAGE/Coomassie Blue staining.

The same methods were used to express and purify Hu-Asp2(b).

**Example 12**

**Assay of Hu-Asp2  $\beta$ -secretase activity using peptide substrates**

*$\beta$ -secretase assay*

Recombinant human Asp2(a) prepared in CHO cells and purified as described in Example 11 was used to assay Asp2(a) proteolytic activity directly. Activity assays for Asp2(a) were performed using synthetic peptide substrates containing either the wild-type APP  $\beta$ -secretase site (SEVKM↓DAEFR), the Swedish KM→NL mutation (SEVNL↓DAEFR), or the A $\beta$ 40 and 42  $\gamma$ -secretase sites (RRGGVV↓IA↓TVTVGER). Reactions were performed in 50 mM 2-[N-morpholino]ethane-sulfonate ("Na-MES," pH 5.5) containing 1%  $\beta$ -octylglucoside, 70 mM peptide substrate, and recombinant Asp2(a) (1-5  $\mu$ g protein) for various times at 37°C. The reaction products were quantified by RP-HPLC using a linear gradient from 0-70 B over 30 minutes (A=0.1% TFA in water, B=0.1%TFA/10%water/90%AcCN). The elution profile was monitored by absorbance at 214 nm. In preliminary experiments, the two product peaks which eluted before the intact peptide substrate, were confirmed to have the sequence DAEFR and SEVNL using both Edman sequencing and MADLI-TOF mass spectrometry. Percent hydrolysis of the peptide substrate was calculated by comparing the integrated peak areas for the two product peptides and the starting material derived from the absorbance at 214 nm. The sequence of cleavage/hydrolysis products was confirmed using Edman sequencing and MADLI-TOF mass spectrometry.

The behavior of purified Asp2(a) in the proteolysis assays was consistent with the prior anti-sense studies which indicated that Asp2(a) possesses  $\beta$ -secretase activity. Maximal proteolysis was seen with the Swedish  $\beta$ -secretase peptide, which, after 6 hours, was about 10-fold higher than wild type APP.

The specificity of the protease cleavage reaction was determined by performing the  $\beta$ -secretase assay in the presence of 8  $\mu$ M pepstatin A and the presence of a cocktail of protease inhibitors (10  $\mu$ M leupeptin, 10  $\mu$ M E64, and 5 mM EDTA). Proteolytic activity was insensitive to both the pepstatin and the cocktail, which are inhibitors of cathepsin D (and other aspartyl proteases), serine proteases, cysteinyl proteases, and metalloproteases, respectively.

Hu-Asp2(b) when similarly expressed in CHO cells and purified using identical conditions for extraction with  $\beta$ -octylglucoside and sequential

chromatography over Mono Q and Mono S also cleaves the Swedish  $\beta$ -secretase peptide in proteolysis assays using identical assay conditions.

Collectively, this data establishes that both forms of Asp2 (Hu-Asp2(a) and Hu-Asp2(b)) act directly in cell-free assays to cleave synthetic APP peptides at the  $\beta$ -secretase site, and that the rate of cleavage is greatly increased by the Swedish KM→NL mutation that is associated with Alzheimer's disease.

An alternative  $\beta$ -secretase assay utilizes internally quenched fluorescent substrates to monitor enzyme activity using fluorescence spectroscopy in a single sample or multiwell format. Each reaction contained 50 mM Na-MES (pH 5.5), peptide substrate MCA-EVKMDAEF[K-DNP] (BioSource International) (50  $\mu$ M) and purified Hu-Asp-2 enzyme. These components were equilibrated to 37 °C for various times and the reaction initiated by addition of substrate. Excitation was performed at 330 nm and the reaction kinetics were monitored by measuring the fluorescence emission at 390 nm. To detect compounds that modulate Hu-Asp-2 activity, the test compounds were added during the preincubation phase of the reaction and the kinetics of the reaction monitored as described above. Activators are scored as compounds that increase the rate of appearance of fluorescence while inhibitors decrease the rate of appearance of fluorescence.

### Example 13

#### Demonstration that Asp1 processes APP at the $\alpha$ -secretase site

Increased expression of an  $\alpha$ -secretase candidate gene in human cells would be expected to increase basal release of sAPP $\alpha$  and to decrease release of A $\beta$  peptides. This the effect was observed when full length human Asp1 is co-expressed with APP in HEK293 cells. The experiment utilized the APP 695 amino acid isoform which had been modified by the addition of a pair of lysine residues to the C-terminus (APP-KK). The C-terminal di-lysine motif increases the intracellular half-life of glycosylated APP and consequently the production of both sAPP $\alpha$  and A $\beta$ . As shown in Table 5, cotransfection of HEK293 cells with APP-KK with Asp1 increased the production of sAPP $\alpha$  by 3.5 fold (p<0.001) and decreased the production of A $\beta$ 40 by 2.8 fold. Thus, Asp1 acts directly or indirectly to facilitate constitutive  $\alpha$ -secretase

cleavage and this effect is competitive with the amyloidogenic processing of APP to A $\beta$  peptides. This implies that mutations or genetic polymorphisms in Asp1 may affect A $\beta$  production by affecting the balance between the competing pathways for constitutive  $\alpha$ -secretase cleavage and A $\beta$  peptide production.

**Table 5.**

**Asp1 stimulates basal release of sAPP $\alpha$  from HEK293 cells after cotransfection with APP-KK.**

Transfection	sAPP $\alpha$ $\mu$ g/ml	Fold Increase	A $\beta$ 40 pg/ml	Fold Decrease
Asp1	3.5 + 1.1	+3.5	113 + 7	-2.8
pcDNA	1.0 + 0.2		321 + 18	

Specific methods used were as follows. The full length Asp1 cDNA was cloned into the vector pcDNA3.1/hygro+(Invitrogen) for transfection studies as previously described (Yan *et al.*, (1999) *Nature* 402: 533-537). The APP-KK cDNA was cloned into the vector pIRES (Clontech) also as previously described. HEK293 cells were transfected with expression constructs using the Lipofectamine Plus reagent from Gibco/BRL. Cells were seeded in 24 well tissue culture plates to a density of 70-80% confluence. Four wells per plate were transfected with 2  $\mu$ g DNA (3:1, APP:Asp1 or empty pcDNA3.1/hygro+ vector), 8  $\mu$ l Plus reagent, and 4  $\mu$ l Lipofectamine in OptiMEM. OptiMEM was added to a total volume of 1 ml, distributed 200  $\mu$ l per well and incubated 3 hours. Care was taken to hold constant the ratios of the two plasmids used for cotransfection as well as the total amount of DNA used in the transfection. The transfection media was replaced with DMEM supplemented with 10% FBS and NaPyruvate, with antibiotic/antimycotic and the cells were incubated under normal conditions (37°, 5% CO<sub>2</sub>) for 48 hours. The conditioned media were removed to polypropylene tubes and stored at -80°C until assayed for the content of sAPP $\alpha$  or A $\beta$ 40/A $\beta$ 42 by enzyme-linked immunosorbent assay (EIA) as described above in Example 6. The A $\beta$  EIA followed the protocol of Pirttila *et al.* (*Neuro. Lett.* (1999) 249: 21-4) using the 6E10 monoclonal antibody (Senetek) as a capture antibody and biotinylated rabbit antiserum 162 or 165 (New

York State Institute for Basic Research, Staten Island, NY) for detection of A $\beta$ 40 and A $\beta$ 42, respectively. The 6E10 antibody recognizes residues 1-16 of the A $\beta$  peptide. The sAPP $\alpha$  EIA used LN27 antibody as a capture antibody and biotinylated 6E10 for detection as described previously (Yan *et al.*, (1999) *supra.*). The LN27 antibody recognized the first 20 amino acids of the human APP peptide.

Increased  $\alpha$ -secretase activity and concomitant reduction of A $\beta$  production *in vivo* represents an effect that may be desirable for the prevention, treatment (*e.g.*, to show the progression of), or cure of Alzheimer's disease. Thus, the activities demonstrated in this example provide an indication that modulators of Asp1 activity, that achieve the same effects *in vivo*, will have utility for Alzheimer's disease therapy. Screening methods for such modulators are contemplated as an aspect of the invention.

#### **Example 14**

##### **Expression of Pre-pro-Hu-Asp1 and Derivatives in Insect Cells**

###### ***Expression of hu-Asp-1TM(His)<sup>6</sup> by baculovirus infection.***

The coding sequence of pre-pro-Hu-Asp1 was engineered for production as a soluble, secreted form by insect cells. PCR primers were designed to (1) delete the predicted transmembrane domain and cytoplasmic tail of Asp1 and (2) to introduce a Kozak consensus sequence for efficient translational initiation. The primers sequences were as follows: sense CGCTTTAAGCTTGCCACCATGGGCGCA CTGGCCCGGGCG (SEQ ID NO: 74) and antisense CGCTTTCTCGAGCTAA TGGTGATGGTGATGGTGCCACAAAATGGGCTCGCTCAAAGA (SEQ ID NO: 75) which replaced the deleted C-terminal transmembrane and cytoplasmic domains with a hexahistidine purification tag.

PCR reactions were carried out with 100 ng of full length Asp1 pcDNA 3.1 hygro+ construct, 200 M NTPs, 300 nM of each primer, 1x reaction buffer containing 2 mM MgSO<sub>4</sub>, and 5 units of Pwo I DNA polymerase (Roche Biochemicals). The reactions were cycled under the following conditions: 94°C for 5 minutes followed by 15 cycles of 94°C for 30 seconds and 72°C for 30 seconds, and then a final extension

reaction at 72°C for 10 minutes. The predicted amino acid sequence of this PCR generated derivative (denoted as Asp-1ΔTM(His)<sub>6</sub>) is set out as SEQ ID NO: 66.

The reaction product was digested to completion with HindIII-XhoI and ligated into the expression vector pIB (Invitrogen) to yield the pIB/Asp-1ΔTM(His)<sub>6</sub> construct. Creation of recombinant baculovirus and infection of sf9 insect cells was performed using standard methods known in the art. Sf9 cells were transfected with either the pIB vector alone or the pIB/Asp-1ΔTM(His)<sub>6</sub> construct utilizing Insectin Plus reagent (Invitrogen) according to the manufacturer's instructions. After the transfection, the cells were cultured in High Five serum-free media (Invitrogen) for 3 days. Subsequently, the conditioned medium was harvested and subjected to Western blot analysis. This analysis revealed specific expression and secretion of immunoreactive Asp-1ΔTM(His)<sub>6</sub> polypeptide into the extracellular medium. The secreted proteins were detected on the Western blot with either the India probe (Pierce Chemicals) specific for the hexahistidine sequence tag or using a rabbit polyclonal antiserum. The polyclonal antisera (denoted as UP-199) was generated by injecting rabbits with recombinant Asp-1ΔTM(His)<sub>6</sub> (SEQ ID NO: 66). This recombinant peptide was prepared by heterologous expression in *E.coli*. The UP-199 antibody recognizes the processed form of Asp-1ΔTM.

Direct analysis with the polyclonal antiserum (UP-199) revealed an immunoreactive band of the expected molecular weight (50 kDa) only in pIB/Asp-1ΔTM(His)<sub>6</sub> transfected cells. This signal was significantly enhanced in concentrated conditioned medium. A similar pattern was obtained using the India probe. No signal was detected in conditioned medium derived from mock-transfected cells using either UP-199 antisera or the India probe.

Based on this result, transient and stable transfections of the pIB/Asp-1ΔTM(His)<sub>6</sub> construct in sf9 insect cells were carried out as described above. Four days post transient transfection, the culture medium was collected to provide material for further characterization. In parallel, sf9 cells were stably transfected with the pIB/Asp-1ΔTM(His)<sub>6</sub> construct and cultured in High Five serum-free medium (Invitrogen) supplemented with 50 µg/ml blasticidin for approximately 2 weeks.



After blasticidin selection, the resistant pool of cells was expanded to provide a stable source of conditioned medium for Asp-1 $\Delta$ TM(His)<sub>6</sub> purification.

*Purification of recombinant Asp-1 $\Delta$ TM(His)<sub>6</sub>*

Conditioned media, from either transient or stably transfected sf9 cells, were concentrated approximately 10-fold using a stirred cell concentrator equipped with a 30,000 MWCO membrane (Spectrum Medical Industries). This concentrate was then subjected to ammonium sulfate precipitation to further concentrate the sample and provide partial purification. Material precipitating between 0-40% saturation was discarded and the resulting supernatant was brought to 80% saturation. Western blot analysis of the various ammonium sulfate precipitated fractions revealed that the majority of the immunoreactive material was contained within the 40-80% ammonium sulfate pellet. As a result, this material was subjected to further purification.

The 40-80% ammonium sulfate pellet was redissolved in approximately 1/20 the original volume of Ni+-NTA loading buffer (25 mM Tris-HCl (pH 8.5)/0.5 M NaCl/10 mM imidazole). Subsequently, the sample was applied to a Ni+-NTA column previously equilibrated in Ni+-NTA buffer. Following sample application, the column was washed with starting buffer (25 mM Tris-HCl (pH 8.5)/ 0.5 M NaCl/ 20 mM imidazole) until the A<sup>280nm</sup> of the column effluent returned to zero. After washing, the bound recombinant protein was eluted off the column with a linear gradient of Ni+-NTA buffer containing increasing concentrations (10 mM, 50 mM, 100 mM, 250 mM, and 500 mM) of imidazole. The elution profile was monitored by Western blot analysis using the UP-199 antiserum. Immunoreactive Asp-1 $\Delta$ TM(His)<sub>6</sub> was detected in the column load and eluted at 50 mM imidazole. NuPAGE gel analysis of the 50 mM imidazole fraction demonstrated a purity of Asp-1 $\Delta$ TM(His)<sub>6</sub> of approximately 50%, therefore further purification was required.

The positive fractions, eluted off the Ni+-NTA column, were then pooled (denoted as post-IMAC pool), concentrated using a YM30 membrane (Amicon), and dialyzed with 25 mM Tris-HCl (pH 8.0). The dialyzed post-IMAC pool was fractionated by MonoQ anion exchange chromatography (Amersham-Pharmacia Biotech) gradient elution containing increasing concentrations (0 -0.5 M) of NaCl

(Buffer A: 25 mM Tris-HCl (pH 8.0) and Buffer B: 25 mM Tris-HCl (pH 8.0)/ 0.5 M NaCl). The elution profile was determined by Western blot analysis which indicated immunoreactive fractions as those displaying immunoreactivity with the UP-199 antisera. NuPAGE gel analysis with silver staining demonstrated that the material prepared in this manner was >90% pure. The immunoactive fractions eluted off the MonoQ anion exchange column were pooled, dialyzed with 25 mM HEPES-Na<sup>+</sup> (pH 8.0), and stored at 4°C until further analysis.

*Acid-activation of recombinant Asp-1 ΔTM(His)<sub>6</sub>*

Recombinant Asp-1 ΔTM(His)<sub>6</sub> migrated with an apparent molecular weight of 50 kD. Direct N-terminal sequence analysis carried out by automated Edman degradation for 20 cycles revealed a unique sequence beginning at Glu<sup>3</sup> (SEQ ID NO: 67), confirming the identity of the recombinant protein. Computer assisted prediction of the signal peptidase cleavage site indicated that the pro-form should initiate at Ala<sup>1</sup>, suggesting either an unusual processing site by the signal peptidase during secretion or an additional processing step that removes an additional two amino acid residues.

To investigate the mechanism of pro-Asp-1 ΔTM(His)<sub>6</sub> activation, aliquots of the purified protein were incubated in various acidic environments with pH values ranging from 3.0-8.0 at 37°C for 2 hours. Subsequently, the recombinant proteins were analyzed by Western blot. A faster migrating polypeptide species was detected after incubation at pH values of 4.0, 4.5 and 5.0. The polypeptide migration was unaltered after incubation in environments which were either more acidic (pH 3.0 and 3.5) or more basic (pH 6.0, 7.0, and 8.0). Sequence analysis of this faster migrating species revealed that it initiated exclusively at Ala<sup>43</sup>, consistent with removal of a 42 amino acid residue segment of the pro-peptide that was induced by treatment of the pro-enzyme at pH 4.5. The predicted amino acid sequence of the acid processed form of Asp-1 ΔTM(His)<sub>6</sub> is set out as SEQ ID NO: 68.

To purify the acid-activated form of Asp-1 ΔTM(His)<sub>6</sub>, the Asp-1 ΔTM(His)<sub>6</sub> post-IMAC pool (generated as described above) was dialyzed to pH 4.5 and then subjected to affinity chromatography on either pepstatin A agarose or sulfolink-PHA-292593E. Following sample application, the column was washed with

25 mM NaOAc (pH 4.5) and eluted with 50 mM Na-BO<sub>3</sub> (pH 9.5). The positive fractions eluted off the columns were dialyzed with 25 mM Hepes-Na (pH 7.5) overnight at 4°C which resulted in quantitative conversion of the pro-enzyme to the acid-processed form (SEQ ID NO: 68) described above. Western blot analysis of the elution profile revealed quantitative retention of immunoreactive Asp-1ΔTM(His)<sub>6</sub> on both affinity resins as evidenced by the lack of Asp-1ΔTM(His)<sub>6</sub> in the unbound fraction as detected by UP-199 immunoreactivity on a Western blot. Step elution 50 mM NaBO<sub>3</sub> at pH 9.5 resulted in elution of immunoreactive Asp-1 ΔTM(His)<sub>6</sub>, with variable recovery.

Comparison of the properties of the recombinant soluble catalytic domain of Asp1 with the properties determined for Asp2 (see Example 10) revealed a number of significant differences. Processing of the pre-pro forms of either enzyme is distinct, with Asp1 undergoing efficient processing by the signal peptidase and additional processing to remove two additional amino acid residues from the N-terminus. Further processing of the pro-form of Asp1 was not detected in neutral pH. In contrast, recombinant Asp2 produced, under similar conditions, yields an equimolar mixture of the pro-form and a processed form that has 24 amino acid residues of the pro-segment removed.

Another distinction between the processing of these two enzymes involves processing initiated by acid-treatment. Systematic analysis of acid-induced processing of pro-Asp2 revealed that the purified polypeptide did not self-process. In contrast, acid dependent processing of pro-Asp1 was readily demonstrated (as described above). Alignment of the self-processing site in Asp1 with the processing site in Asp2 revealed that these two enzymes are processed at the same position, which is a different method of processing as compared with that of other known human aspartyl proteases.

In addition to providing valuable information about Asp1 activity, the discovery of a site of apparent autocatalytic processing of Asp1 provides an indication of a peptide sequences (surrounding Ala<sup>43</sup>) that could be useful for performing

screening assays to identify modulators of Asp1 activity. This idea is explored in greater detail in Example 15.

### Example 15

#### Development of an enzymatic assay for Asp-1 $\Delta$ TM(His)<sub>6</sub>

The relationship between Asp1 and APP processing was explored by determining if APP  $\alpha$ -secretase, APP  $\beta$ -secretase, or APP  $\gamma$ -secretase peptide substrates were cleaved by recombinant Asp-1 $\Delta$ TM(His)<sub>6</sub>. These peptide substrates included the  $\alpha$ -secretase specific substrates A $\beta$ <sub>10-20</sub> and A $\beta$ <sub>12-28</sub>, the  $\beta$ -secretase specific substrates PHA-95812E (SEVKMDAEFR; SEQ ID NO: 64) and PHA-247574E (SEVNLDAEFR; SEQ ID NO: 63), and  $\gamma$ -secretase specific substrate PHA-179111E (RRGGVVIATVIVGER; SEQ ID NO: 76). Each reaction consisted of incubating a peptide substrate (100 nM) with recombinant Asp-1 $\Delta$ TM(His)<sub>6</sub> for 15 hours at pH 4.5 at 37°C. Reaction products were quantified by RP-HPLC at A<sup>214 nm</sup>. The elution profiles for Asp-1 $\Delta$ TM(His)<sub>6</sub> were compared to those obtained from parallel Asp1 experiments. The identity of the cleavage products was determined by MADLI-TOF mass spectrometry. Table 6 summarizes the Asp1 substrates and indicates the cleavage site.

**Table 6**

#### Substrate Preferences of Asp-1 $\Delta$ TM

<u>P4</u>	<u>P3</u>	<u>P2</u>	<u>P1</u>		<u>P1'</u>	<u>P2'</u>	<u>P3'</u>	<u>P4'</u>		SEQ ID NO:
G	L	A	L	↓	A	L	E	P	Self Activation	69
E	V	K	M	↓	D	A	E	F	$\beta$ -Secretase, WT	70
E	V	N	L	↓	D	A	E	F	$\beta$ -Secretase, Sw	71
L	V	F	F	↓	A	E	D	V	A $\beta$ <sub>12-28</sub> ( $\alpha$ -Secretase)	72
K	L	V	F	↓	F	A	E	D	A $\beta$ <sub>12-28</sub> ( $\alpha$ -Secretase)	73

The peptides in Table 6 are described using the nomenclature by Schechter and Berger (*Biochem. Biophys. Res. Commun.* 27:157(1967) and *Biochem. Biophys. Res. Commun.* 32:898 (1968)), in which the amino acid residues in the peptide substrate that undergo the cleavage are defined as  $P_1 \dots P_n$  toward the N-terminus and  $P_1' \dots P_n'$  toward the C-terminus. Therefore, the scissile bond is between the  $P_1$  and the  $P_1'$  residue of the peptide subunits and is denoted herein throughout with a hyphen between the  $P_1$  and the  $P_1'$ .

Digestion of the  $\alpha$ -secretase substrate ( $A_{12-28}$ ) revealed two Asp1 cleavage sites. The major product was cleaved at  $\text{Phe}^{20}\downarrow\text{Ala}^{21}$  and the minor product was cleaved at  $\text{Phe}^{19}\downarrow\text{Phe}^{20}$  (referring to the numbering convention in the APP A $\beta$ ) peptide. Analysis of the cleavage products obtained from the  $\beta$ -secretase peptide substrates revealed that both the wild-type (PHA-95812E) and the Swedish mutation (PHA-247574E) substrates were hydrolyzed exclusively at the  $\beta$ -secretase site. Also, the relative rates of Asp-1-dependent hydrolysis of the  $\beta$ -secretase peptide substrate containing the Swedish mutation was cleaved at least 10-times faster than the corresponding wild-type peptide. Conversion of the  $\gamma$ -secretase peptide substrate was not detected under these reaction conditions.

Measurement of the cleavage of the  $\alpha$ -secretase and  $\beta$ -secretase substrates can also be carried out with substrates comprising detectable labels such as radioactive, enzymatic, chemiluminescent or fluorescent labels. For example, the peptide substrates could comprise internally quenched labels that result in increased detectability after cleavage of the peptide substrates due to separation of the labels upon cleavage. The peptide substrates can be modified to have attached a paired fluorophore and quencher such as 7-amino-4-methyl coumarin and dinitrophenol, respectively.

This example illustrates the  $\alpha$ -secretase and  $\beta$ -secretase activity exhibited by Asp-1, confirming the APP processing activity of Asp1 indicating, e.g., in Examples 7 and 13. The substrates described herein may be used in combination with recombinant Asp1 to measure Asp1 proteolytic activity at the  $\alpha$ -secretase and  $\beta$ -secretase

processing sites. These substrates are useful in screening assays for identification of modulators of Asp1 proteolytic activity.

In particular, production of A $\beta$  species through the processing of APP at  $\beta$ - and  $\gamma$ -secretase sites may play a central role in Alzheimer's disease pathogenesis, and processing at the  $\alpha$ -secretase site may have a protective role and may prevent A $\beta$  production. Thus, a therapeutic and/or prophylactic indication exists for molecules that can increase Asp1  $\alpha$ -secretase activity and/or decrease Asp1  $\beta$ -secretase activity *in vivo*. The present invention includes screening assays for such modulators, and the foregoing substrate peptides are useful in such assays.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the invention. The entire disclosure of all publications cited herein are hereby incorporated by reference.

## SEQUENCE LISTING

<110> Pharmacia & Upjohn

<120> ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES THEREFOR

<130> 28341/6280PCP

<140>

<141>

<150> 60/169,232

<151> 1999-12-06

<150> 09/416,901

<151> 1999-10-13

<150> 60/155,493

<151> 1999-09-23

<150> 09/404,133

<151> 1999-09-23

<150> PCT/US99/20881

<151> 1999-09-23

<150> 60/101,594

<151> 1998-09-24

<160> 76

<170> PatentIn Ver. 2.0

<210> 1

<211> 1804

<212> DNA

<213> Homo sapiens

<400> 1

```

atggggcgcac tggcccgggc gctgctgctg cctctgctgg ccagtggtt cctgcgcgcc 60
gccccggagc tggccccgc gcccttcacg ctgccccctc ggggtggcgc ggccacgaac 120
cgcgtagtgt cgccaccccc gggaccggg acccctgcc agcggcacgc cgacggcttg 180
gcgctcgccc tggagcctgc cctggcgctc ccgcggggc ccgccaactt cttggccatg 240
gtagacaacc tgcaggggga ctctggccgc ggctactacc tggagatgct gatcgggacc 300
ccccgcaga agctacagat tctcgttgac actggaagca gtaactttgc cgtggcagga 360
acccgcact cctacataga cacgtacttt gacacagaga ggtctagcac ataccgctcc 420
aagggctttg acgtcacagt gaagtacaca caaggaagct ggacgggctt cgttggggaa 480
gacctcgtca ccatcccaa aggcctcaat acttcttttc ttgtcaacat tgccactatt 540
tttgaatcag agaatttctt tttgctggg attaatgga atggaatact tggcctagct 600
tatgccacac ttgccaagcc atcaagttct ctggagacct tcttcgactc cctggtgaca 660
caagcaaaca tccccaacgt tttctccatg cagatgtgtg gagccggctt gcccgttgct 720
ggatctggga ccaacggagg tagtctgtg ttgggtgga ttgaaccaag tttgtataaa 780
ggagacatct ggtatacccc tattaaggaa gagtggtagt accagataga aattctgaaa 840
ttggaaattg gagggcaaaag ccttaatctg gactgcagag agtataacgc agacaaggcc 900
atcgtggaca gtggcaccac gctgctgcgc ctgccccaga aggtgtttga tgcggtggtg 960
gaagctgtgg ccgcgcgcatc tctgattcca gaattctctg atgggtttctg gactgggtcc 1020
cagctggcgt gctggacgaa ttcggaaaca ccttgggtctt acttcctaa aatctccatc 1080
tacctgagag atgagaactc cagcaggtca ttccgtatca caatcctgcc tcagctttac 1140
attcagccca tgatgggggc cggcctgaat tatgaatgtt accgattcgg catttcccca 1200
tccacaaatg cgctggtgat cggtgccacg gtgatggagg gcttctacgt catcttcgac 1260

```

```

agagcccaga agagggtggg cttcgcagcg agcccctgtg cagaaattgc aggtgctgca 1320
gtgtctgaaa tttccgggcc tttctcaaca gaggatgtag ccagcaactg tgtccccgct 1380
cagtctttga gcgagcccat tttgtggatt gtgtcctatg cgctcatgag cgtctgtgga 1440
gccatcctcc ttgtcttaat cgtcctgctg ctgctgccgt tccggtgtca gcgtcgcccc 1500
cgtgaccctg aggtcgtcaa tgatgagtcc tctctgggtc gacatcgctg gaaatgaata 1560
gccaggcctg acctcaagca accatgaact cagctattaa gaaaatcaca tttccagggc 1620
agcagccggg atcgatggtg gcgctttctc ctgtgcccac ccgtcttcaa tctctgttct 1680
gctcccagat gccttctaga ttcactgtct tttgattctt gattttcaag ctttcaaadc 1740
ctccctactt ccaagaaaaa taattaaata aaaaacttca ttctaaacca aaaaaaaaaa 1800
aaaa                                              1804

```

<210> 2

<211> 518

<212> PRT

<213> Homo sapiens

<400> 2

```

Met Gly Ala Leu Ala Arg Ala Leu Leu Leu Pro Leu Leu Ala Gln Trp
  1              5              10              15

Leu Leu Arg Ala Ala Pro Glu Leu Ala Pro Ala Pro Phe Thr Leu Pro
      20              25              30

Leu Arg Val Ala Ala Ala Thr Asn Arg Val Val Ala Pro Thr Pro Gly
      35              40              45

Pro Gly Thr Pro Ala Glu Arg His Ala Asp Gly Leu Ala Leu Ala Leu
      50              55              60

Glu Pro Ala Leu Ala Ser Pro Ala Gly Ala Ala Asn Phe Leu Ala Met
      65              70              75              80

Val Asp Asn Leu Gln Gly Asp Ser Gly Arg Gly Tyr Tyr Leu Glu Met
      85              90              95

Leu Ile Gly Thr Pro Pro Gln Lys Leu Gln Ile Leu Val Asp Thr Gly
      100             105             110

Ser Ser Asn Phe Ala Val Ala Gly Thr Pro His Ser Tyr Ile Asp Thr
      115             120             125

Tyr Phe Asp Thr Glu Arg Ser Ser Thr Tyr Arg Ser Lys Gly Phe Asp
      130             135             140

Val Thr Val Lys Tyr Thr Gln Gly Ser Trp Thr Gly Phe Val Gly Glu
      145             150             155             160

Asp Leu Val Thr Ile Pro Lys Gly Phe Asn Thr Ser Phe Leu Val Asn
      165             170             175

Ile Ala Thr Ile Phe Glu Ser Glu Asn Phe Phe Leu Pro Gly Ile Lys
      180             185             190

Trp Asn Gly Ile Leu Gly Leu Ala Tyr Ala Thr Leu Ala Lys Pro Ser
      195             200             205

Ser Ser Leu Glu Thr Phe Phe Asp Ser Leu Val Thr Gln Ala Asn Ile
      210             215             220

Pro Asn Val Phe Ser Met Gln Met Cys Gly Ala Gly Leu Pro Val Ala
      225             230             235             240

```



Gly Ser Gly Thr Asn Gly Gly Ser Leu Val Leu Gly Gly Ile Glu Pro  
 245 250 255  
 Ser Leu Tyr Lys Gly Asp Ile Trp Tyr Thr Pro Ile Lys Glu Glu Trp  
 260 265 270  
 Tyr Tyr Gln Ile Glu Ile Leu Lys Leu Glu Ile Gly Gly Gln Ser Leu  
 275 280 285  
 Asn Leu Asp Cys Arg Glu Tyr Asn Ala Asp Lys Ala Ile Val Asp Ser  
 290 295 300  
 Gly Thr Thr Leu Leu Arg Leu Pro Gln Lys Val Phe Asp Ala Val Val  
 305 310 315 320  
 Glu Ala Val Ala Arg Ala Ser Leu Ile Pro Glu Phe Ser Asp Gly Phe  
 325 330 335  
 Trp Thr Gly Ser Gln Leu Ala Cys Trp Thr Asn Ser Glu Thr Pro Trp  
 340 345 350  
 Ser Tyr Phe Pro Lys Ile Ser Ile Tyr Leu Arg Asp Glu Asn Ser Ser  
 355 360 365  
 Arg Ser Phe Arg Ile Thr Ile Leu Pro Gln Leu Tyr Ile Gln Pro Met  
 370 375 380  
 Met Gly Ala Gly Leu Asn Tyr Glu Cys Tyr Arg Phe Gly Ile Ser Pro  
 385 390 395 400  
 Ser Thr Asn Ala Leu Val Ile Gly Ala Thr Val Met Glu Gly Phe Tyr  
 405 410 415  
 Val Ile Phe Asp Arg Ala Gln Lys Arg Val Gly Phe Ala Ala Ser Pro  
 420 425 430  
 Cys Ala Glu Ile Ala Gly Ala Ala Val Ser Glu Ile Ser Gly Pro Phe  
 435 440 445  
 Ser Thr Glu Asp Val Ala Ser Asn Cys Val Pro Ala Gln Ser Leu Ser  
 450 455 460  
 Glu Pro Ile Leu Trp Ile Val Ser Tyr Ala Leu Met Ser Val Cys Gly  
 465 470 475 480  
 Ala Ile Leu Leu Val Leu Ile Val Leu Leu Leu Leu Pro Phe Arg Cys  
 485 490 495  
 Gln Arg Arg Pro Arg Asp Pro Glu Val Val Asn Asp Glu Ser Ser Leu  
 500 505 510  
 Val Arg His Arg Trp Lys  
 515

&lt;210&gt; 3

&lt;211&gt; 2070

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 3

```

atggcccaag ccctgccctg gctcctgctg tggatgggcg cgggagtgct gcctgccac 60
ggcaccacagc acggcatccg gctgcccctg cgcagcggcc tggggggcg cccctgggg 120
ctgcggtctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
gtgggcagcc ccccgagac gctcaacatc ctgggtggata caggcagcag taactttgca 300
gtgggtgctg ccccccaccc ctctcctgcat cgctactacc agaggcagct gtccagcaca 360
taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420
ctgggcaccg acctggtaag catcccccat ggcccccaacg tcaactgtgcg tgccaacatt 480
gctgccatca ctgaatcaga caagtctctc atcaacggct ccaactggga aggcattcctg 540
gggtggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 600
ctggtaaagc agaccacgt tcccaacctc ttctccctgc acctttgtgg tgctgcttc 660
cccccaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 720
gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 780
gaggatcatca ttgtcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 840
tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 900
gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 960
ggtttctggc taggagagca gctgggtgtc tggcaagcag gcaccacccc ttggaacatt 1020
ttcccagtca tctactcta cctaattgggt gaggttacca accagtcctt ccgcatcacc 1080
atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtacca agacgactgt 1140
tacaagtttg ccatctcaca gtcattccag ggcactgtta tgggagctgt tatcatggag 1200
ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1260
catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1320
gaagactgtg gctacaacat tccacagaca gatgagtcaa ccctcatgac catagcctat 1380
gtcatggctg ccatctgcgc cctcttcctg ctgccactct gcctcatggt gtgtcagtgg 1440
cgctgcctcc gctgcctgcg ccagcagcat gatgactttg ctgatgacat ctccctgctg 1500
aagtgaggag gcccatgggc agaagataga gattccctg gaccacacct ccgtggttca 1560
ctttggtcac aagtaggaga cacagatggc acctgtggcc agagcacctc aggaccctcc 1620
ccaccacca aatgcctctg ccttgatgga gaaggaaaag gctggcaagg tgggttccag 1680
ggactgtacc ttaggaaac agaaaagaga agaaagaagc actctgctgg cggaataact 1740
cttggtcacc tcaaatttaa gtcgggaaat tctgctgctt gaaacttcag ccctgaacct 1800
ttgtccacca ttctttaaa ttctccaacc caaagtattc ttcttttctt agtttcagaa 1860
gtactggcat cacacgcagg ttaccttggc gtgtgtccct gtggtaccct ggcagagaag 1920
agaccaagct tgttccctg ctggccaaag tcagtaggag aggatgcaca gtttgctatt 1980
tgctttagag acaggactg tataaacaag cctaacattg gtgcaaagat tgcctcttga 2040
attaaaaaaa aaaaaaaaaa aaaaaaaaaa

```

&lt;210&gt; 4

&lt;211&gt; 501

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

```

Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val
  1              5              10              15

Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser
  20              25              30

Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
  35              40              45

Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
  50              55              60

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
  65              70              75              80

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
  85              90              95

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr

```

100						105						110					
Tyr	Gln	Arg	Gln	Leu	Ser	Ser	Thr	Tyr	Arg	Asp	Leu	Arg	Lys	Gly	Val		
115						120						125					
Tyr	Val	Pro	Tyr	Thr	Gln	Gly	Lys	Trp	Glu	Gly	Glu	Leu	Gly	Thr	Asp		
130						135						140					
Leu	Val	Ser	Ile	Pro	His	Gly	Pro	Asn	Val	Thr	Val	Arg	Ala	Asn	Ile		
145						150						155					
Ala	Ala	Ile	Thr	Glu	Ser	Asp	Lys	Phe	Phe	Ile	Asn	Gly	Ser	Asn	Trp		
165						170						175					
Glu	Gly	Ile	Leu	Gly	Leu	Ala	Tyr	Ala	Glu	Ile	Ala	Arg	Pro	Asp	Asp		
180						185						190					
Ser	Leu	Glu	Pro	Phe	Phe	Asp	Ser	Leu	Val	Lys	Gln	Thr	His	Val	Pro		
195						200						205					
Asn	Leu	Phe	Ser	Leu	His	Leu	Cys	Gly	Ala	Gly	Phe	Pro	Leu	Asn	Gln		
210						215						220					
Ser	Glu	Val	Leu	Ala	Ser	Val	Gly	Gly	Ser	Met	Ile	Ile	Gly	Gly	Ile		
225						230						235					
Asp	His	Ser	Leu	Tyr	Thr	Gly	Ser	Leu	Trp	Tyr	Thr	Pro	Ile	Arg	Arg		
245						250						255					
Glu	Trp	Tyr	Tyr	Glu	Val	Ile	Ile	Val	Arg	Val	Glu	Ile	Asn	Gly	Gln		
260						265						270					
Asp	Leu	Lys	Met	Asp	Cys	Lys	Glu	Tyr	Asn	Tyr	Asp	Lys	Ser	Ile	Val		
275						280						285					
Asp	Ser	Gly	Thr	Thr	Asn	Leu	Arg	Leu	Pro	Lys	Lys	Val	Phe	Glu	Ala		
290						295						300					
Ala	Val	Lys	Ser	Ile	Lys	Ala	Ala	Ser	Ser	Thr	Glu	Lys	Phe	Pro	Asp		
305						310						315					
Gly	Phe	Trp	Leu	Gly	Glu	Gln	Leu	Val	Cys	Trp	Gln	Ala	Gly	Thr	Thr		
325						330						335					
Pro	Trp	Asn	Ile	Phe	Pro	Val	Ile	Ser	Leu	Tyr	Leu	Met	Gly	Glu	Val		
340						345						350					
Thr	Asn	Gln	Ser	Phe	Arg	Ile	Thr	Ile	Leu	Pro	Gln	Gln	Tyr	Leu	Arg		
355						360						365					
Pro	Val	Glu	Asp	Val	Ala	Thr	Ser	Gln	Asp	Asp	Cys	Tyr	Lys	Phe	Ala		
370						375						380					
Ile	Ser	Gln	Ser	Ser	Thr	Gly	Thr	Val	Met	Gly	Ala	Val	Ile	Met	Glu		
385						390						395					
Gly	Phe	Tyr	Val	Val	Phe	Asp	Arg	Ala	Arg	Lys	Arg	Ile	Gly	Phe	Ala		
405						410						415					
Val	Ser	Ala	Cys	His	Val	His	Asp	Glu	Phe	Arg	Thr	Ala	Ala	Val	Glu		
420						425						430					

Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro  
 435 440 445

Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala  
 450 455 460

Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp  
 465 470 475 480

Arg Cys Leu Arg Cys Leu Arg Gln Gln His Asp Asp Phe Ala Asp Asp  
 485 490 495

Ile Ser Leu Leu Lys  
 500

<210> 5

<211> 1977

<212> DNA

<213> Homo sapiens

<400> 5

```

atggcccaag cctgcccctg gctcctgctg tggatgggag cgggagtgtc gcctgcccac 60
ggcaccacag acggcatccg gctgcccctg cgcagcggcc tggggggcgc cccctggggg 120
ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
gtgggcagcc ccccgagac gctcaacatc ctggtggata caggcagcag taactttgca 300
gtgggtgctg ccccccaccc ctctctgcat cgctactacc agaggcagct gtccagcaca 360
taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420
ctgggcaccg acctggttaag catcccccat ggccccaacg tcaactgtgc tgccaacatt 480
gctgcatca ctgaatcaga caagtctctc atcaacggct ccaactggga aggcacacct 540
gggctggcct atgctgagat tgccaggctt tgtgtgtgtg gcttccccct caaccagtct 600
gaagtgtgtg cctctgtcgg agggagcatg atcattggag gtatcgacca ctcgctgtac 660
acaggcagtc tctggtatcc acccatccgg cgggagtggg attatgaggt gatcattgtg 720
cgggtggaga tcaatggaca ggatctgaaa atggactgca aggagtacaa ctatgacaag 780
agcattgttg acagtggcac caccaacctt cgtttgccc aaaaagtgtt tgaagctgca 840
gtcaaatcca tcaaggcagc ctctccacg gagaagttcc ctgatgggtt ctggctagga 900
gagcagctgg tgtgctggca agcaggcacc accccttggg acattttccc agtcatctca 960
ctctacctaa tgggtgaggt taccaaccag tccttcgca tcaccatcct tccgcagcaa 1020
tacctgcggc cagtggaaaga tgtggccacg tcccaagacg actgttacia gtttgccatc 1080
tcacagtcac ccacgggcac tgttatggga gctgttatca tggagggcct ctacgttgtc 1140
tttgatcggg cccgaaaacg aattggcttt gctgtcagcg cttgccatgt gcacgatgag 1200
ttcaggacgg cagcgttggg agggcctttt gtcaccttgg acatggaaga ctgtggctac 1260
aacattccac agacagatga gtcaaccctc atgaccatag cctatgtcat ggctgccatc 1320
tgcgccctct tcatgctgcc actctgcctc atgggtgtgtc agtggcgtg cctccgctgc 1380
ctgcgccagc agcatgatga ctttgcctgac gacatctccc tgctgaagtg aggaggccca 1440
tgggcagaag atagagattc ccttgaccca cactccgtg gttcactttg gtcacaagta 1500
ggagacacag atggcacctg tggccagagc acctcaggac cctccccacc caccaaatgc 1560
ctctgccttg atggagaagg aaaaggctgg caagggtggg tccagggact gtacctgtag 1620
gaaacagaaa agagaagaaa gaagcactct gctggcgagg atactcttgg tcacctcaaa 1680
tttaagtctg gaaattctgc tgcttgaaac ttcagccctg aacctttgtc caccattcct 1740
ttaaattctc caaccaaaag tattcttctt ttcttagttt cagaagtact ggcacacac 1800
gcagggtacc ttggcgtgtg tccctgtggt accctggcag agaagagacc aagcttgttt 1860
ccctgctggc caaagtcagt aggagaggat gcacagtttg ctatttgctt tagagacagg 1920
gactgtataa acaagcctaa cattggtgca aagattgcct cttgaaaaaa aaaaaaa 1977

```

<210> 6

<211> 476

<212> PRT

<213> Homo sapiens

&lt;400&gt; 6

Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val  
 1 5 10 15  
 Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser  
 20 25 30  
 Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp  
 35 40 45  
 Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val  
 50 55 60  
 Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr  
 65 70 75 80  
 Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser  
 85 90 95  
 Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr  
 100 105 110  
 Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val  
 115 120 125  
 Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp  
 130 135 140  
 Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile  
 145 150 155 160  
 Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp  
 165 170 175  
 Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Leu Cys Gly  
 180 185 190  
 Ala Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly  
 195 200 205  
 Ser Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu  
 210 215 220  
 Trp Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val  
 225 230 235 240  
 Arg Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr  
 245 250 255  
 Asn Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu  
 260 265 270  
 Pro Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser  
 275 280 285  
 Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val  
 290 295 300  
 Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser  
 305 310 315 320  
 Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile

325	330	335
Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln		
340	345	350
Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val		
355	360	365
Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala		
370	375	380
Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu		
385	390	395
Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu		
405	410	415
Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser Thr Leu Met Thr		
420	425	430
Ile Ala Tyr Val Met Ala Ala Ile Cys Ala Leu Phe Met Leu Pro Leu		
435	440	445
Cys Leu Met Val Cys Gln Trp Arg Cys Leu Arg Cys Leu Arg Gln Gln		
450	455	460
His Asp Asp Phe Ala Asp Asp Ile Ser Leu Leu Lys		
465	470	475

&lt;210&gt; 7

&lt;211&gt; 2043

&lt;212&gt; DNA

&lt;213&gt; Mus musculus

&lt;400&gt; 7

```

atggccccag cgctgcactg gctcctgcta tgggtgggct cggaatgct gcctgcccag 60
ggaaccatc tcggcatccg gctgcccctt cgcagcggcc tggcagggcc acccctgggc 120
ctgaggctgc cccgggagac tgacgaggaa tcggaggagc ctggccggag aggcagcttt 180
gtggagatgg tggacaacct gaggggaaag tccggccagg gctactatgt ggagatgacc 240
gtaggcagcc cccacacagac gctcaacatc ctggtggaca cgggcagtag taactttgca 300
gtgggggctg cccacacccc ttctctgcat cgctactacc agaggcagct gtccagcaca 360
tatcgagacc tccgaaaggg tgtgtatgtg ccctacaccc agggcaagtg ggagggggaa 420
ctgggcaccg acctggtgag catccctcat ggcccaacg tcaactgtcg tgccaacatt 480
gctgccatca ctgaatcgga caagttcttc atcaatggtt ccaactggga gggcatccta 540
gggctggcct atgctgagat tgccaggccc gacgactctt tggagccctt ctttgactcc 600
ctggtgaagc agaccacat tcccaacatc tttccctgc agctctgtgg cgctggcttc 660
cccctcaacc agaccgaggc actggcctcg gtgggaggga gcatgatcat tgggtggtatc 720
gaccactcgc tatacacggg cagtctctgg tacacacca tccggcggga gtggtattat 780
gaagtgatca ttgtacgtgt ggaaatcaat ggtcaagatc tcaagatgga ctgcaaggag 840
tacaactacg acaagagcat tgtggacagt gggaccacca accttcgctt gcccaagaaa 900
gtatttgaag ctgccgtcaa gtccatcaag gcagcctcct cgacggagaa gttcccggat 960
ggcttttggc taggggagca gctggtgtgc tggcaagcag gcacgacccc ttggaacatt 1020
ttcccagtc tttcacttta cctcatgggt gaagtcacca atcagtcctt ccgcatcacc 1080
atccttcttc agcaatacct acggccggtg gaggacgtgg ccacgtcca agacgactgt 1140
tacaagtctg ctgtctcaca gtcattccag ggcaactgta tgggagccgt catcatggaa 1200
ggtttctatg tcgtcttcga tcgagcccga aagcgaattg gctttgctgt cagcgcttgc 1260
catgtgcacg atgagttcag gacggcggca gtggaaggtc cgtttggttac ggcagacatg 1320
gaagactgtg gctacaacat tcccagaca gatgagtcac cacttatgac catagcctat 1380
gtcatggcgg ccatctgcgc cctcttcatg ttgccactct gectcatggt atgtcagtg 1440
cgctgcctgc gttgcctgcg ccaccagcac gatgactttg ctgatgacat ctccctgctc 1500

```

```

aagtaaggag gctcgtgggc agatgatgga gacgccctg gaccacatct ggggtggtcc 1560
ctttggtcac atgagttgga gctatggatg gtacctgtgg ccagagcacc tcaggaccct 1620
caccaacctg ccaatgcttc tggcgtgaca gaacagagaa atcaggcaag ctggattaca 1680
gggcttgcac ctgtaggaca caggagaggg aagggaagcag cgttctggtg gcaggaatat 1740
ccttaggcac cacaaacttg agttggaaat tttgctgctt gaagcttcag ccctgaccct 1800
ctgcccagca tcctttagag tctccaacct aaagtattct ttatgtcctt ccagaagtac 1860
tggcgtcata ctcaggctac cgggcatgtg tccctgtggt accctggcag agaaagggcc 1920
aatctcattc cctgctggcc aaagtcagca gaagaagggtg aagtttgcca gttgctttag 1980
tgatagggac tgcagactca agcctacact ggtacaaaga ctgcgtcttg agataaacia 2040
gaa 2043

```

<210> 8

<211> 501

<212> PRT

<213> Mus musculus

<400> 8

```

Met Ala Pro Ala Leu His Trp Leu Leu Leu Trp Val Gly Ser Gly Met
  1              5              10              15

```

```

Leu Pro Ala Gln Gly Thr His Leu Gly Ile Arg Leu Pro Leu Arg Ser
              20              25              30

```

```

Gly Leu Ala Gly Pro Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp
              35              40              45

```

```

Glu Glu Ser Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val
              50              55              60

```

```

Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr
              65              70              75              80

```

```

Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser
              85              90              95

```

```

Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr
              100              105              110

```

```

Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val
              115              120              125

```

```

Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp
              130              135              140

```

```

Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile
              145              150              155              160

```

```

Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp
              165              170              175

```

```

Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp
              180              185              190

```

```

Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Ile Pro
              195              200              205

```

```

Asn Ile Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln
              210              215              220

```

```

Thr Glu Ala Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile
              225              230              235              240

```

Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg  
 245 250 255  
 Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln  
 260 265 270  
 Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val  
 275 280 285  
 Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala  
 290 295 300  
 Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp  
 305 310 315 320  
 Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr  
 325 330 335  
 Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val  
 340 345 350  
 Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg  
 355 360 365  
 Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala  
 370 375 380  
 Val Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu  
 385 390 395 400  
 Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala  
 405 410 415  
 Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu  
 420 425 430  
 Gly Pro Phe Val Thr Ala Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro  
 435 440 445  
 Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val Met Ala Ala  
 450 455 460  
 Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val Cys Gln Trp  
 465 470 475 480  
 Arg Cys Leu Arg Cys Leu Arg His Gln His Asp Asp Phe Ala Asp Asp  
 485 490 495  
 Ile Ser Leu Leu Lys  
 500

&lt;210&gt; 9

&lt;211&gt; 2088

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 9

atgctgcccc gtttggcact gctcctgctg gccgcctgga cggtcggggc gctggaggta 60  
 cccactgatg gtaatgctgg cctgctggct gaacccaga ttgcatgtt ctgtggcaga 120



```

ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcattcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagtgtgg 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttactggt cacaccgtcg ccaaagagac atgcagttag 480
aagagtacca acttgcattg ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggaggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgccca ccaccaccac caccaccaca 840
gagtcctgtg aagaggtggg tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaaagc ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaagag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtggg agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttctcctc cgccctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggttg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tcctgtctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctcttccccg tgaatggaga gttagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggag gtgttgtcat agcgacagtg 1920
atcgtcatca ccttgggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgcccgtgt caccacagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaactag 2088

```

<210> 10

<211> 695

<212> PRT

<213> Homo sapiens

<400> 10

```

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
  1              5              10              15

```

```

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
      20              25              30

```

```

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
      35              40              45

```

```

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
      50              55              60

```

```

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
      65              70              75              80

```

```

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
      85              90              95

```

```

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
      100              105              110

```

```

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu

```

115	120	125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys 130 135 140		
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu 145 150 155 160		
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile 165 170 175		
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu 180 185 190		
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val 195 200 205		
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys 210 215 220		
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu 225 230 235 240		
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu 245 250 255		
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile 260 265 270		
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg 275 280 285		
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu 290 295 300		
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys 305 310 315 320		
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg 325 330 335		
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp 340 345 350		
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu 355 360 365		
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala 370 375 380		
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn 385 390 395 400		
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe 405 410 415		
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His 420 425 430		
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala 435 440 445		

Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu  
 450 455 460  
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala  
 465 470 475 480  
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn  
 485 490 495  
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser  
 500 505 510  
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr  
 515 520 525  
 Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln  
 530 535 540  
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn  
 545 550 555 560  
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr  
 565 570 575  
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser  
 580 585 590  
 Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val  
 595 600 605  
 His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys  
 610 615 620  
 Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val  
 625 630 635 640  
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile  
 645 650 655  
 His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg  
 660 665 670  
 His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys  
 675 680 685  
 Phe Phe Glu Gln Met Gln Asn  
 690 695

<210> 11  
 <211> 2088  
 <212> DNA  
 <213> Homo sapiens

<400> 11  
 atgctgcccc gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60  
 cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120  
 ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaa 180  
 acctgcattg ataccaagga aggcacctg cagtattgcc aagaagtcta ccctgaactg 240  
 cagatcacca atgtggtaga agccaaccaa ccagtgaacca tccagaactg gtgcaagcgg 300

```

ggccgcaagc agtgcaagac ccattccccac tttgtgattc cctaccgctg cttagtgtggt 360
gagtttgttaa gtgatgccct tctcggttcc gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttctactg cacaccgtcg ccaaagagac atgcagtggag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcgaggagg atgactcgga tgtctggttg ggccggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtctgttg aagaggtggt tctagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccattttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctgggtggag 1140
acacacatgg ccagagtgga agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacacc taaagcattt cgagcatgtg 1320
cgcatggttg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcatgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaacta ttcagatgac 1500
gtcttggtcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctcttctccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctggtgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgatcat agcgacagtg 1920
atcgtcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgctgt caccacagag gagegccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaacct caagttcttt gagcagatgc agaactag 2088

```

<210> 12  
 <211> 695  
 <212> PRT  
 <213> Homo sapiens

<400> 12  
 Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg  
 1 5 10 15  
 Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro  
 20 25 30  
 Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln  
 35 40 45  
 Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp  
 50 55 60  
 Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu  
 65 70 75 80  
 Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn  
 85 90 95  
 Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val  
 100 105 110  
 Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu  
 115 120 125  
 Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys

130		135		140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu				
145		150		155 160
Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile				
	165		170	175
Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu				
	180		185	190
Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val				
	195		200	205
Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys				
	210		215	220
Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu				
	225		230	235 240
Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu				
		245	250	255
Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile				
	260		265	270
Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg				
	275		280	285
Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu				
	290		295	300
Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys				
	305		310	315 320
Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg				
		325	330	335
Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp				
		340	345	350
Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu				
	355		360	365
Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala				
	370		375	380
Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn				
	385		390	395 400
Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe				
		405	410	415
Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His				
	420		425	430
Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala				
	435		440	445
Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu				
	450		455	460

Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala  
465 470 475 480

Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn  
485 490 495

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser  
500 505 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr  
515 520 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln  
530 535 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn  
545 550 555 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr  
565 570 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser  
580 585 590

Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val  
595 600 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys  
610 615 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val  
625 630 635 640

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile  
645 650 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg  
660 665 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys  
675 680 685

Phe Phe Glu Gln Met Gln Asn  
690 695

<210> 13

<211> 2088

<212> DNA

<213> Homo sapiens

<400> 13

atgctgccccg gtttggcact gtcctgctg gccgcctgga cggctcgggc gctggaggta 60  
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120  
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaa 180  
acctgcattg ataccaagga aggcacctg cagtattgcc aagaagtcta ccctgaactg 240  
cagatcacca atgtggtaga agccaaccaa ccagtgaacca tccagaactg gtgcaagcgg 300  
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggg 360  
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420  
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagttag 480

```

aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcgaggaggg atgactcgga tgtctgggtg ggaggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgccca ccaccaccac caccaccaca 840
gagtctgttg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctgggtggag 1140
acacacatgg ccagagtggg agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatggttg atcccaagaa agccgctcag atccggctcc aggttatgac acacctccgt 1380
gtgatttatg agcgcataaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaaag agcaaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccctcccg tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atgggtggcg gtgttgatc atgcagctg 1920
atcttcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg agcccgctgt caccacagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaactag 2088

```

<210> 14

<211> 695

<212> PRT

<213> Homo sapiens

<400> 14

```

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1             5             10             15

```

```

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
          20             25             30

```

```

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
          35             40             45

```

```

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
          50             55             60

```

```

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
          65             70             75             80

```

```

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
          85             90             95

```

```

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
          100            105            110

```

```

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
          115            120            125

```

```

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
          130            135            140

```

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu  
 145 150 155 160  
 Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile  
 165 170 175  
 Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu  
 180 185 190  
 Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val  
 195 200 205  
 Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys  
 210 215 220  
 Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu  
 225 230 235 240  
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu  
 245 250 255  
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile  
 260 265 270  
 Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg  
 275 280 285  
 Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu  
 290 295 300  
 Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys  
 305 310 315 320  
 Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg  
 325 330 335  
 Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp  
 340 345 350  
 Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu  
 355 360 365  
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala  
 370 375 380  
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn  
 385 390 395 400  
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe  
 405 410 415  
 Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His  
 420 425 430  
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala  
 435 440 445  
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu  
 450 455 460  
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala



465		470		475		480
Glu Glu Ile Gln Asp	Glu Val Asp	Glu Leu Leu Gln Lys	Glu Gln Asn			
485		490	495			
Tyr Ser Asp Asp Val	Leu Ala Asn Met	Ile Ser Glu Pro Arg	Ile Ser			
500		505	510			
Tyr Gly Asn Asp Ala	Leu Met Pro Ser	Leu Thr Glu Thr	Lys Thr Thr			
515		520	525			
Val Glu Leu Leu Pro	Val Asn Gly Glu Phe	Ser Leu Asp Asp	Leu Gln			
530		535	540			
Pro Trp His Ser Phe	Gly Ala Asp Ser	Val Pro Ala Asn Thr	Glu Asn			
545		550	555			560
Glu Val Glu Pro Val	Asp Ala Arg Pro	Ala Ala Asp Arg	Gly Leu Thr			
	565	570	575			
Thr Arg Pro Gly Ser	Gly Leu Thr Asn	Ile Lys Thr Glu	Glu Ile Ser			
	580	585	590			
Glu Val Lys Met Asp	Ala Glu Phe Arg	His Asp Ser Gly	Tyr Glu Val			
	595	600	605			
His His Gln Lys Leu	Val Phe Phe Ala	Glu Asp Val Gly	Ser Asn Lys			
	610	615	620			
Gly Ala Ile Ile Gly	Leu Met Val Gly	Gly Val Val Ile	Ala Thr Val			
	630	635	640			
Ile Phe Ile Thr Leu	Val Met Leu Lys	Lys Lys Gln Tyr	Thr Ser Ile			
	645	650	655			
His His Gly Val Val	Glu Val Asp Ala	Ala Val Thr Pro	Glu Glu Arg			
	660	665	670			
His Leu Ser Lys Met	Gln Gln Asn Gly	Tyr Glu Asn Pro	Thr Tyr Lys			
	675	680	685			
Phe Phe Glu Gln Met	Gln Asn					
	690	695				

&lt;210&gt; 15

&lt;211&gt; 2094

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 15

```

atgctgcccg gtttggcact gtcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaa 180
acctgcattg ataccaagga aggcattcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtaccca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttagt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagttag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600

```

```

gcgaggagg atgactcggg tgtctggtgg ggcggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagctctgtg aagaggtggt tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggagttat ccagcatttc 1080
caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctggtggag 1140
acacacatgg ccagagtggg agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200
tacatcaccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaac gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcattggtg atcccaagaa agccgctcag atccggtccc aggttatgac acacctccgt 1380
gtgatttatg agcgcataaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaaag agcaaaacta ttcagatgac 1500
gtcttgccca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560
tctttgaccg aaacgaaaac caccgtggag ctccttcccg tgaatggaga gttagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgcccgtg gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgatc agcgacagtg 1920
atcgatcatc ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggagggtg acgccgctgt caccacagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaacaagaa gtag 2094

```

<210> 16

<211> 697

<212> PRT

<213> Homo sapiens

<400> 16

```

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1          5          10         15
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
          20          25         30
Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
          35          40         45
Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
          50          55         60
Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
          65          70         75         80
Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
          85          90         95
Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
          100         105        110
Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
          115        120        125
Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
          130        135        140
Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
          145        150        155        160

```

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile  
 165 170 175  
 Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu  
 180 185 190  
 Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val  
 195 200 205  
 Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys  
 210 215 220  
 Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu  
 225 230 235 240  
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu  
 245 250 255  
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile  
 260 265 270  
 Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg  
 275 280 285  
 Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu  
 290 295 300  
 Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys  
 305 310 315 320  
 Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg  
 325 330 335  
 Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp  
 340 345 350  
 Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu  
 355 360 365  
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala  
 370 375 380  
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn  
 385 390 395 400  
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe  
 405 410 415  
 Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His  
 420 425 430  
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala  
 435 440 445  
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu  
 450 455 460  
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala  
 465 470 475 480  
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn

485					490					495					
Tyr	Ser	Asp	Asp	Val	Leu	Ala	Asn	Met	Ile	Ser	Glu	Pro	Arg	Ile	Ser
			500					505					510		
Tyr	Gly	Asn	Asp	Ala	Leu	Met	Pro	Ser	Leu	Thr	Glu	Thr	Lys	Thr	Thr
		515					520					525			
Val	Glu	Leu	Leu	Pro	Val	Asn	Gly	Glu	Phe	Ser	Leu	Asp	Asp	Leu	Gln
	530					535					540				
Pro	Trp	His	Ser	Phe	Gly	Ala	Asp	Ser	Val	Pro	Ala	Asn	Thr	Glu	Asn
545					550					555					560
Glu	Val	Glu	Pro	Val	Asp	Ala	Arg	Pro	Ala	Ala	Asp	Arg	Gly	Leu	Thr
				565					570					575	
Thr	Arg	Pro	Gly	Ser	Gly	Leu	Thr	Asn	Ile	Lys	Thr	Glu	Glu	Ile	Ser
			580					585					590		
Glu	Val	Lys	Met	Asp	Ala	Glu	Phe	Arg	His	Asp	Ser	Gly	Tyr	Glu	Val
		595					600					605			
His	His	Gln	Lys	Leu	Val	Phe	Phe	Ala	Glu	Asp	Val	Gly	Ser	Asn	Lys
	610					615					620				
Gly	Ala	Ile	Ile	Gly	Leu	Met	Val	Gly	Gly	Val	Val	Ile	Ala	Thr	Val
625					630					635					640
Ile	Val	Ile	Thr	Leu	Val	Met	Leu	Lys	Lys	Lys	Gln	Tyr	Thr	Ser	Ile
				645					650					655	
His	His	Gly	Val	Val	Glu	Val	Asp	Ala	Ala	Val	Thr	Pro	Glu	Glu	Arg
			660					665					670		
His	Leu	Ser	Lys	Met	Gln	Gln	Asn	Gly	Tyr	Glu	Asn	Pro	Thr	Tyr	Lys
	675						680					685			
Phe	Phe	Glu	Gln	Met	Gln	Asn	Lys	Lys							
	690					695									

&lt;210&gt; 17

&lt;211&gt; 2094

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 17

```

atgctgcccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcattctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttggt 360
gagtttgtaa gtgatgccct tctcgttctt gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcaactg cacaccgtcg ccaaagagac atgcagtgag 480
aagagtacca acttgcatga ctacggcatg ttgtgcccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggttg gccggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780

```

```

ccctacgaag aagccacaga gagaaccacc agcattgccca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggg tgcagttcct acaacagcag ccagtacccc tgatgccgtt 900
gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960
gagaggcttg aggccaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020
gcagaacgtc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080
caggagaaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctgggtggag 1140
acacacatgg ccagagtggg agccatgctc aatgaccgcc gccgcctggc cctgggagaac 1200
tacatcacccg ctctgcaggc tgttcctcct cggcctcgtc acgtgttcaa tatgctaaag 1260
aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320
cgcatgggtg atcccaagaa agccgctcag atccgggtccc aggttatgac acacctccgt 1380
gtgatattatg agcgcagtaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440
gaggagattc aggatgaagt tgatgagctg cttcagaaag agcaaaaacta ttcagatgac 1500
gtcttggcca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgccca 1560
tctttgaccg aaacgaaaac caccgtggag ctcttcccgc tgaatggaga gttcagcctg 1620
gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680
gaagttgagc ctgttgatgc ccgccctgct gccgaccgag gactgaccac tcgaccaggt 1740
tctgggttga caaatatcaa gacggaggag atctctgaag tgaatctgga tgcagaattc 1800
cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860
ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920
atcgtcatca ccttgggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980
gtggaggttg acgccgtgt caccacagag gagcgccacc tgtccaagat gcagcagaac 2040
ggctacgaaa atccaaccta caagttcttt gagcagatgc agaacaagaa gtag 2094

```

<210> 18

<211> 697

<212> PRT

<213> Homo sapiens

<400> 18

```

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1              5              10             15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
      20              25              30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
      35              40              45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
      50              55              60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
      65              70              75             80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
      85              90              95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
      100             105             110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
      115             120             125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
      130             135             140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
      145             150             155             160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
      165             170             175

```

Asp	Lys	Phe	Arg	Gly	Val	Glu	Phe	Val	Cys	Cys	Pro	Leu	Ala	Glu	Glu	180	185	190	
Ser	Asp	Asn	Val	Asp	Ser	Ala	Asp	Ala	Glu	Glu	Asp	Asp	Ser	Asp	Val	195	200	205	
Trp	Trp	Gly	Gly	Ala	Asp	Thr	Asp	Tyr	Ala	Asp	Gly	Ser	Glu	Asp	Lys	210	215	220	
Val	Val	Glu	Val	Ala	Glu	Glu	Glu	Glu	Val	Ala	Glu	Val	Glu	Glu	Glu	225	230	235	240
Glu	Ala	Asp	Asp	Asp	Glu	Asp	Asp	Glu	Asp	Gly	Asp	Glu	Val	Glu	Glu	245	250	255	
Glu	Ala	Glu	Glu	Pro	Tyr	Glu	Glu	Ala	Thr	Glu	Arg	Thr	Thr	Ser	Ile	260	265	270	
Ala	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Glu	Ser	Val	Glu	Glu	Val	Val	Arg	275	280	285	
Val	Pro	Thr	Thr	Ala	Ala	Ser	Thr	Pro	Asp	Ala	Val	Asp	Lys	Tyr	Leu	290	295	300	
Glu	Thr	Pro	Gly	Asp	Glu	Asn	Glu	His	Ala	His	Phe	Gln	Lys	Ala	Lys	305	310	315	320
Glu	Arg	Leu	Glu	Ala	Lys	His	Arg	Glu	Arg	Met	Ser	Gln	Val	Met	Arg	325	330	335	
Glu	Trp	Glu	Glu	Ala	Glu	Arg	Gln	Ala	Lys	Asn	Leu	Pro	Lys	Ala	Asp	340	345	350	
Lys	Lys	Ala	Val	Ile	Gln	His	Phe	Gln	Glu	Lys	Val	Glu	Ser	Leu	Glu	355	360	365	
Gln	Glu	Ala	Ala	Asn	Glu	Arg	Gln	Gln	Leu	Val	Glu	Thr	His	Met	Ala	370	375	380	
Arg	Val	Glu	Ala	Met	Leu	Asn	Asp	Arg	Arg	Arg	Leu	Ala	Leu	Glu	Asn	385	390	395	400
Tyr	Ile	Thr	Ala	Leu	Gln	Ala	Val	Pro	Pro	Arg	Pro	Arg	His	Val	Phe	405	410	415	
Asn	Met	Leu	Lys	Lys	Tyr	Val	Arg	Ala	Glu	Gln	Lys	Asp	Arg	Gln	His	420	425	430	
Thr	Leu	Lys	His	Phe	Glu	His	Val	Arg	Met	Val	Asp	Pro	Lys	Lys	Ala	435	440	445	
Ala	Gln	Ile	Arg	Ser	Gln	Val	Met	Thr	His	Leu	Arg	Val	Ile	Tyr	Glu	450	455	460	
Arg	Met	Asn	Gln	Ser	Leu	Ser	Leu	Leu	Tyr	Asn	Val	Pro	Ala	Val	Ala	465	470	475	480
Glu	Glu	Ile	Gln	Asp	Glu	Val	Asp	Glu	Leu	Leu	Gln	Lys	Glu	Gln	Asn	485	490	495	

Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser  
 500 505 510  
 Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr  
 515 520 525  
 Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln  
 530 535 540  
 Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn  
 545 550 555 560  
 Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr  
 565 570 575  
 Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser  
 580 585 590  
 Glu Val Asn Leu Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val  
 595 600 605  
 His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys  
 610 615 620  
 Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val  
 625 630 635 640  
 Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile  
 645 650 655  
 His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg  
 660 665 670  
 His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys  
 675 680 685  
 Phe Phe Glu Gln Met Gln Asn Lys Lys  
 690 695

&lt;210&gt; 19

&lt;211&gt; 2094

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 19

atgctgccccg gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60  
 cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120  
 ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaa 180  
 acctgcattg ataccaagga aggcacctct cagtattgcc aagaagtcta ccctgaactg 240  
 cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300  
 ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagtgtgt 360  
 gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420  
 atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtgg 480  
 aagagtacca acttgcata gaacggcatg ttgctgacct gcggaattga caagttccga 540  
 ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgtgat 600  
 gcggaggagg atgactcggg tgtctggtgg ggcggagcag acacagacta tgcagatggg 660  
 agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720  
 gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780  
 ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840  
 gagtctgtgg aagaggtgtg tcgagttcct acaacagcag ccagtacccc tgatgccgtt 900  
 gacaagtatc tcgagacacc tggggatgag aatgaacatg cccatttcca gaaagccaaa 960

gagaggcttg aggccaaagca ccgagagaga atgtcccagg tcatgagaga atgggaagag 1020  
 gcagaacgctc aagcaaagaa cttgcctaaa gctgataaga aggcagttat ccagcatttc 1080  
 caggagaaag tggaatcttt ggaacaggaa gcagccaacg agagacagca gctgggtggag 1140  
 acacacatgg ccagagtggg agccatgctc aatgaccgcc gccgcctggc cctggagaac 1200  
 tacatcaccg ctctgcaggc tgttctctct cggcctcgtc acgtgttcaa tatgctaaag 1260  
 aagtatgtcc gcgcagaaca gaaggacaga cagcacaccc taaagcattt cgagcatgtg 1320  
 cgcattggtg atcccaagaa agccgctcag atccgggtccc aggttatgac acacctccgt 1380  
 gtgatttatg agcgcattgaa tcagtctctc tccctgctct acaacgtgcc tgcagtggcc 1440  
 gaggagattc aggatgaagt tgatgagctg cttcagaaaag agcaaaacta ttcagatgac 1500  
 gtcttgccca acatgattag tgaaccaagg atcagttacg gaaacgatgc tctcatgcca 1560  
 tctttggacc aaacgaaaac caccgtggag ctcttcccg tgaatggaga gttcagcctg 1620  
 gacgatctcc agccgtggca ttcttttggg gctgactctg tgccagccaa cacagaaaac 1680  
 gaagttgagc ctgttgatgc ccgcctgct gccgaccgag gactgaccac tcgaccaggt 1740  
 tctgggttga caaatatcaa gacggaggag atctctgaag tgaagatgga tgcagaattc 1800  
 cgacatgact caggatatga agttcatcat caaaaattgg tgttctttgc agaagatgtg 1860  
 ggttcaaaca aaggtgcaat cattggactc atggtgggcg gtgttgtcat agcgacagtg 1920  
 atcttcatca ccttggtgat gctgaagaag aaacagtaca catccattca tcatggtgtg 1980  
 gtggagggtg acgccgctgt caccacagag gagcgccacc tgtccaagat gcagcagaac 2040  
 ggctacgaaa atccaaccta caagttcttt gagcagatgc agaacaagaa gtag 2094

<210> 20

<211> 697

<212> PRT

<213> Homo sapiens

<400> 20

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg  
 1 5 10 15

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro  
 20 25 30

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln  
 35 40 45

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp  
 50 55 60

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu  
 65 70 75 80

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn  
 85 90 95

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val  
 100 105 110

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu  
 115 120 125

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys  
 130 135 140

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu  
 145 150 155 160

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile  
 165 170 175

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu  
 180 185 190



Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val  
 195 200 205  
 Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys  
 210 215 220  
 Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu  
 225 230 235 240  
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu  
 245 250 255  
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile  
 260 265 270  
 Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg  
 275 280 285  
 Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu  
 290 295 300  
 Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys  
 305 310 315 320  
 Glu Arg Leu Glu Ala Lys His Arg Glu Arg Met Ser Gln Val Met Arg  
 325 330 335  
 Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp  
 340 345 350  
 Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu  
 355 360 365  
 Gln Glu Ala Ala Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala  
 370 375 380  
 Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn  
 385 390 395 400  
 Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe  
 405 410 415  
 Asn Met Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His  
 420 425 430  
 Thr Leu Lys His Phe Glu His Val Arg Met Val Asp Pro Lys Lys Ala  
 435 440 445  
 Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu  
 450 455 460  
 Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala  
 465 470 475 480  
 Glu Glu Ile Gln Asp Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn  
 485 490 495  
 Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser  
 500 505 510

Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr  
515 520 525

Val Glu Leu Leu Pro Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln  
530 535 540

Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn  
545 550 555 560

Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr  
565 570 575

Thr Arg Pro Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser  
580 585 590

Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val  
595 600 605

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys  
610 615 620

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val  
625 630 635 640

Ile Phe Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile  
645 650 655

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg  
660 665 670

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys  
675 680 685

Phe Phe Glu Gln Met Gln Asn Lys Lys  
690 695

<210> 21

<211> 1341

<212> DNA

<213> Homo sapiens

<400> 21

```

atggctagca tgactggtgg acagcaaagt ggtcgcggtat ccaccagca cggcatccgg 60
ctgcccctgc gcagcggcct ggggggcgcc cccctggggc tgcggctgcc ccgggagacc 120
gacgaagagc ccgaggagcc cggccggagg ggcagctttg tggagatggt ggacaacctg 180
aggggcaagt cggggcaggg ctactacgtg gagatgaccg tgggcagccc cccgcagacg 240
ctcaacatcc tgggtgatac aggcagcagt aactttgcag tgggtgctgc cccccacccc 300
ttcctgcatc gctactacca gaggcagctg tccagcacat accgggacct ccggaagggt 360
gtgtatgtgc cctacacca gggcaagtgg gaaggggagc tgggcaccga cctggttaagc 420
atcccccatg gcccacagt cactgtgcgt gccaacattg ctgccatcac tgaatcagac 480
aagttcttca tcaacggctc caactgggaa ggcattcctg ggctggccta tgetgagatt 540
gccaggcctg acgactccct ggagcctttc tttgactctc tggtaaagca gaccacggtt 600
cccaacctct tctccctgca cctttgtggt gctggcttcc ccctcaacca gtctgaagtg 660
ctggcctctg tcggagggag catgatcatt ggaggtatcg accactcgct gtacacaggc 720
agtctctggt atacacccat ccggcgggag tggattatg aggtcatcat tgtgcggtg 780
gagatcaatg gacaggatct gaaaatggac tgcaaggagt acaactatga caagagcatt 840
gtggacagtg gcaccaccaa ccttcgtttg cccaagaaag tgtttgaagc tgcagtcaaa 900
tccatcaagg cagcctcctc cacggagaag ttccctgatg gtttctggct aggagagcag 960
ctggtgtgct ggcaagcagg caccaccctt tggaacattt tcccagtcac ctactctac 1020
ctaattgggtg aggttaccaa ccagtccttc cgcattacca tccttcgca gcaatacctg 1080

```

cggccagtgg aagatgtggc cacgtcccaa gacgactgtt acaagtttgc catctcacag 1140  
 tcatccacgg gcaactgttat gggagctgtt atcatggagg gcttctacgt tgtctttgat 1200  
 cgggcccga aacgaattgg ctttgctgtc agcgcttgcc atgtgcacga tgagttcagg 1260  
 acggcagcgg tggaaggccc tttgtgcacc ttggacatgg aagactgtgg ctacaacatt 1320  
 ccacagacag atgagtcatg a 1341

<210> 22

<211> 446

<212> PRT

<213> Homo sapiens

<400> 22

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Thr Gln  
 1 5 10 15

His Gly Ile Arg Leu Pro Leu Arg Ser Gly Leu Gly Gly Ala Pro Leu  
 20 25 30

Gly Leu Arg Leu Pro Arg Glu Thr Asp Glu Glu Pro Glu Glu Pro Gly  
 35 40 45

Arg Arg Gly Ser Phe Val Glu Met Val Asp Asn Leu Arg Gly Lys Ser  
 50 55 60

Gly Gln Gly Tyr Tyr Val Glu Met Thr Val Gly Ser Pro Pro Gln Thr  
 65 70 75 80

Leu Asn Ile Leu Val Asp Thr Gly Ser Ser Asn Phe Ala Val Gly Ala  
 85 90 95

Ala Pro His Pro Phe Leu His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser  
 100 105 110

Thr Tyr Arg Asp Leu Arg Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly  
 115 120 125

Lys Trp Glu Gly Glu Leu Gly Thr Asp Leu Val Ser Ile Pro His Gly  
 130 135 140

Pro Asn Val Thr Val Arg Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp  
 145 150 155 160

Lys Phe Phe Ile Asn Gly Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala  
 165 170 175

Tyr Ala Glu Ile Ala Arg Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp  
 180 185 190

Ser Leu Val Lys Gln Thr His Val Pro Asn Leu Phe Ser Leu His Leu  
 195 200 205

Cys Gly Ala Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val  
 210 215 220

Gly Gly Ser Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly  
 225 230 235 240

Ser Leu Trp Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile  
 245 250 255

Ile Val Arg Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys  
 260 265 270

Glu Tyr Asn Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu  
275 280 285

Arg Leu Pro Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala  
290 295 300

Ala Ser Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln  
305 310 315 320

Leu Val Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val  
325 330 335

Ile Ser Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile  
340 345 350

Thr Ile Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr  
355 360 365

Ser Gln Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly  
370 375 380

Thr Val Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp  
385 390 395 400

Arg Ala Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His  
405 410 415

Asp Glu Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp  
420 425 430

Met Glu Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser  
435 440 445

<210> 23

<211> 1380

<212> DNA

<213> Homo sapiens

<400> 23

```

atggctagca tgactggtgg acagcaaagt ggtcgcggat cgatgactat ctctgactct 60
ccgcgtgaac aggacggatc caccagcac ggcacccg tgcccctgcg cagcggcctg 120
gggggcgccc ccctggggct gcggctgccc cgggagaccg acgaagagcc cgaggagccc 180
ggccggaggg gcagctttgt ggagatggtg gacaacctga ggggcaagtc ggggcagggc 240
tactacgtgg agatgaccgt gggcagcccc ccgcagacgc tcaacatcct ggtggataca 300
ggcagcagta actttgcagt ggggtgctgcc cccacccct tcctgcatcg ctactaccag 360
aggcagctgt ccagcacata ccgggacctc cggaagggtg tgtatgtgcc ctacaccag 420
ggcaagtggg aaggggagct gggcaccgac ctggtaagca tccccatgg cccaacgctc 480
actgtgcgtg ccaacattgc tgccatcact gaatcagaca agttcttcat caacggctcc 540
aactgggaag gcatectggg gctggcctat gctgagattg ccaggcctga cgactccctg 600
gagcctttct ttgactctct ggtaaagcag acccaggttc ccaacctctt ctccctgcac 660
ctttgtggtg ctggcttccc cctcaaccag tctgaagtgc tggcctctgt cggagggagc 720
atgatcattg gaggtatcga ccaactcgctg tacacaggca gtctctggta tacaccatc 780
cggcgggagt ggtattatga ggtcatcatt gtgcgggtgg agatcaatgg acaggatctg 840
aaaatggact gcaaggagta caactatgac aagagcattg tggacagtgg caccaccaac 900
cttcgtttgc ccaagaaagt gtttgaagct gcagtc aaat ccatcaaggc agcctcctcc 960
acggagaagt tccctgatgg tttctggcta ggagagcagc tgggtgtgctg gcaagcaggc 1020
accacccctt ggaacatttt cccagtcac tcactctacc taatgggtga ggttaccaac 1080
cagtccttcc gcacaccat ccttcgcgag caatacctgc ggccagtga agatgtggcc 1140
acgtcccaag acgactgtta caagtttgcc atctcacagt catccacggg cactgttatg 1200

```

ggagctgtta tcatggaggg cttctacgtt gtctttgata gggcccgaaa acgaattggc 1260  
 ttgtctgtca gcgcttgcca tgtgcacgat gagttcagga cggcagcggg ggaaggccct 1320  
 tttgtcacct tggacatgga agactgtggc tacaacattc cacagacaga tgagtcatga 1380

<210> 24

<211> 459

<212> PRT

<213> Homo sapiens

<400> 24

Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Gly Ser Met Thr  
 1 5 10 15

Ile Ser Asp Ser Pro Arg Glu Gln Asp Gly Ser Thr Gln His Gly Ile  
 20 25 30

Arg Leu Pro Leu Arg Ser Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg  
 35 40 45

Leu Pro Arg Glu Thr Asp Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly  
 50 55 60

Ser Phe Val Glu Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly  
 65 70 75 80

Tyr Tyr Val Glu Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile  
 85 90 95

Leu Val Asp Thr Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His  
 100 105 110

Pro Phe Leu His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg  
 115 120 125

Asp Leu Arg Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu  
 130 135 140

Gly Glu Leu Gly Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val  
 145 150 155 160

Thr Val Arg Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe  
 165 170 175

Ile Asn Gly Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu  
 180 185 190

Ile Ala Arg Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp Ser Leu Val  
 195 200 205

Lys Gln Thr His Val Pro Asn Leu Phe Ser Leu His Leu Cys Gly Ala  
 210 215 220

Gly Phe Pro Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly Ser  
 225 230 235 240

Met Ile Ile Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu Trp  
 245 250 255

Tyr Thr Pro Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg  
 260 265 270

Val Glu Ile Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn

275	280	285
Tyr Asp Lys Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro		
290	295	300
Lys Lys Val Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser Ser		
305	310	315 320
Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val Cys		
	325	330 335
Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser Leu		
	340	345 350
Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu		
	355	360 365
Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln Asp		
	370	375 380
Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val Met		
385	390	395 400
Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala Arg		
	405	410 415
Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu Phe		
	420	425 430
Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu Asp		
	435	440 445
Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser		
450	455	

&lt;210&gt; 25

&lt;211&gt; 1302

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 25

```

atgactcagc atggtattcg tctgccactg cgtagcggtc tgggtggtgc tccactgggt 60
ctgcgtctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 120
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 180
gtgggcagcc ccccgagac gctcaacatc ctggtggata caggcagcag taactttgca 240
gtgggtgctg cccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 300
taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 360
ctgggcaccg acctggtgtaag catcccccat ggcccccaacg tcaactgtcg tgccaacatt 420
gtgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcattcctg 480
gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 540
ctggtaaagc agaccacagt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 600
cccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggtatc 660
gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 720
gaggtcatca ttgtgcgggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 780
tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 840
gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 900
gtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 960
ttcccagtc tctcactcta cctaattgggt gaggttacca accagtcctt ccgcatcacc 1020
atccttccgc agcaatacct ggggccagtg gaagatgtgg ccacgtccca agacgactgt 1080
tacaagtttg ccatctcaca gtcattccag ggcactgtta tgggagctgt tatcatggag 1140

```

```
<210> 26
<211> 433
<212> PRT
<213> Homo sapiens
```

<400> 26																
Met	Thr	Gln	His	Gly	Ile	Arg	Leu	Pro	Leu	Arg	Ser	Gly	Leu	Gly	Gly	
1				5					10					15		
Ala	Pro	Leu	Gly	Leu	Arg	Leu	Pro	Arg	Glu	Thr	Asp	Glu	Glu	Pro	Glu	
			20					25					30			
Glu	Pro	Gly	Arg	Arg	Gly	Ser	Phe	Val	Glu	Met	Val	Asp	Asn	Leu	Arg	
		35					40					45				
Gly	Lys	Ser	Gly	Gln	Gly	Tyr	Tyr	Val	Glu	Met	Thr	Val	Gly	Ser	Pro	
	50					55					60					
Pro	Gln	Thr	Leu	Asn	Ile	Leu	Val	Asp	Thr	Gly	Ser	Ser	Asn	Phe	Ala	
65					70					75					80	
Val	Gly	Ala	Ala	Pro	His	Pro	Phe	Leu	His	Arg	Tyr	Tyr	Gln	Arg	Gln	
				85					90					95		
Leu	Ser	Ser	Thr	Tyr	Arg	Asp	Leu	Arg	Lys	Gly	Val	Tyr	Val	Pro	Tyr	
			100					105					110			
Thr	Gln	Gly	Lys	Trp	Glu	Gly	Glu	Leu	Gly	Thr	Asp	Leu	Val	Ser	Ile	
		115					120					125				
Pro	His	Gly	Pro	Asn	Val	Thr	Val	Arg	Ala	Asn	Ile	Ala	Ala	Ile	Thr	
	130					135					140					
Glu	Ser	Asp	Lys	Phe	Phe	Ile	Asn	Gly	Ser	Asn	Trp	Glu	Gly	Ile	Leu	
145					150					155					160	
Gly	Leu	Ala	Tyr	Ala	Glu	Ile	Ala	Arg	Pro	Asp	Asp	Ser	Leu	Glu	Pro	
				165					170					175		
Phe	Phe	Asp	Ser	Leu	Val	Lys	Gln	Thr	His	Val	Pro	Asn	Leu	Phe	Ser	
			180					185					190			
Leu	His	Leu	Cys	Gly	Ala	Gly	Phe	Pro	Leu	Asn	Gln	Ser	Glu	Val	Leu	
		195					200					205				
Ala	Ser	Val	Gly	Gly	Ser	Met	Ile	Ile	Gly	Gly	Ile	Asp	His	Ser	Leu	
		210				215					220					
Tyr	Thr	Gly	Ser	Leu	Trp	Tyr	Thr	Pro	Ile	Arg	Arg	Glu	Trp	Tyr	Tyr	
225					230					235					240	
Glu	Val	Ile	Ile	Val	Arg	Val	Glu	Ile	Asn	Gly	Gln	Asp	Leu	Lys	Met	
				245					250					255		
Asp	Cys	Lys	Glu	Tyr	Asn	Tyr	Asp	Lys	Ser	Ile	Val	Asp	Ser	Gly	Thr	
			260					265					270			
Thr	Asn	Leu	Arg	Leu	Pro	Lys	Lys	Val	Phe	Glu	Ala	Ala	Val	Lys	Ser	

275	280	285
Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp Gly Phe Trp Leu		
290	295	300
Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr Pro Trp Asn Ile		
305	310	315 320
Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser		
	325	330 335
Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp		
	340	345 350
Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser		
	355	360 365
Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val		
	370	375 380
Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys		
	385	390 395 400
His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val		
	405	410 415
Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu		
	420	425 430

Ser

&lt;210&gt; 27

&lt;211&gt; 1278

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 27

```

atggctagca tgactggtgg acagcaaagt ggctcgcgat cgatgactat ctctgactct 60
ccgctggact ctggtatcga aaccgacgga tcctttgtgg agatggtgga caacctgagg 120
ggcaagtctgg ggcagggcta ctacgtggag atgaccgtgg gcagccccc gcagacgctc 180
aacatctctgg tggatacagg cagcagtaac tttgcagtgg gtgctgcccc ccacccttc 240
ctgcctcgtct actaccagag gcagctgtcc agcacatacc gggacctccg gaaggggtgtg 300
tatgtgccct acaccaggg caagtgggaa ggggagctgg gcaccgacct ggtaagcatc 360
ccccatggcc ccaacgtcac tgtgcgtgcc aacattgtct ccatcactga atcagacaag 420
ttcttcatca acggctccaa ctgggaaggc atcctggggc tggcctatgc tgagattgcc 480
aggcctgacg actccctgga gcctttcttt gactctctgg taaagcagac ccacgttccc 540
aacctcttct cctgcacct ttgtggtgct ggcttcccc tcaaccagtc tgaagtgtctg 600
gcctctgtcg gagggagcat gatcattgga ggtatcgacc actcgctgta cacaggcagt 660
ctctggtata caccatccg gcgggagtgg tattatgagg tcatcattgt gcgggtggag 720
atcaatggac aggatctgaa aatggactgc aaggagtaca actatgacaa gagcattgtg 780
gacagtggca ccaccaacct tcgtttgccc aagaaagtgt ttgaagctgc agtcaaattcc 840
atcaaggcag cctcctccac ggagaagttc cctgatggtt tctggctagg agagcagctg 900
gtgtgctggc aagcaggcac cacccttgg aacattttcc cagtcattct actctacct 960
atgggtgagg ttaccaacca gtccttccgc atcaccatcc ttccgcagca atacctgcgg 1020
ccagtgggaag atgtggccac gtcccaagac gactgttaca agtttgccat ctcacagtca 1080
tccacgggca ctgttatggg agctgttata atggagggct tctacgttgt ctttgatcgg 1140
gcccgaatac gaattggctt tgctgtcagc gcttgccatg tgcacgatga gttcaggacg 1200
gcagcgggtg aaggcccttt tgtcaccttg gacatggaag actgtggcta caacattcca 1260
cagacagatg agtcatga

```



<210> 28  
 <211> 425  
 <212> PRT  
 <213> Homo sapiens

<400> 28

Met	Ala	Ser	Met	Thr	Gly	Gly	Gln	Gln	Met	Gly	Arg	Gly	Ser	Met	Thr
1				5					10					15	
Ile	Ser	Asp	Ser	Pro	Leu	Asp	Ser	Gly	Ile	Glu	Thr	Asp	Gly	Ser	Phe
			20					25					30		
Val	Glu	Met	Val	Asp	Asn	Leu	Arg	Gly	Lys	Ser	Gly	Gln	Gly	Tyr	Tyr
		35					40					45			
Val	Glu	Met	Thr	Val	Gly	Ser	Pro	Pro	Gln	Thr	Leu	Asn	Ile	Leu	Val
		50				55					60				
Asp	Thr	Gly	Ser	Ser	Asn	Phe	Ala	Val	Gly	Ala	Ala	Pro	His	Pro	Phe
	65				70					75					80
Leu	His	Arg	Tyr	Tyr	Gln	Arg	Gln	Leu	Ser	Ser	Thr	Tyr	Arg	Asp	Leu
				85					90					95	
Arg	Lys	Gly	Val	Tyr	Val	Pro	Tyr	Thr	Gln	Gly	Lys	Trp	Glu	Gly	Glu
			100					105					110		
Leu	Gly	Thr	Asp	Leu	Val	Ser	Ile	Pro	His	Gly	Pro	Asn	Val	Thr	Val
		115					120					125			
Arg	Ala	Asn	Ile	Ala	Ala	Ile	Thr	Glu	Ser	Asp	Lys	Phe	Phe	Ile	Asn
		130				135					140				
Gly	Ser	Asn	Trp	Glu	Gly	Ile	Leu	Gly	Leu	Ala	Tyr	Ala	Glu	Ile	Ala
	145				150					155					160
Arg	Pro	Asp	Asp	Ser	Leu	Glu	Pro	Phe	Phe	Asp	Ser	Leu	Val	Lys	Gln
				165					170					175	
Thr	His	Val	Pro	Asn	Leu	Phe	Ser	Leu	His	Leu	Cys	Gly	Ala	Gly	Phe
			180					185					190		
Pro	Leu	Asn	Gln	Ser	Glu	Val	Leu	Ala	Ser	Val	Gly	Gly	Ser	Met	Ile
		195					200					205			
Ile	Gly	Gly	Ile	Asp	His	Ser	Leu	Tyr	Thr	Gly	Ser	Leu	Trp	Tyr	Thr
	210					215						220			
Pro	Ile	Arg	Arg	Glu	Trp	Tyr	Tyr	Glu	Val	Ile	Ile	Val	Arg	Val	Glu
	225				230					235					240
Ile	Asn	Gly	Gln	Asp	Leu	Lys	Met	Asp	Cys	Lys	Glu	Tyr	Asn	Tyr	Asp
				245					250					255	
Lys	Ser	Ile	Val	Asp	Ser	Gly	Thr	Thr	Asn	Leu	Arg	Leu	Pro	Lys	Lys
			260					265					270		
Val	Phe	Glu	Ala	Ala	Val	Lys	Ser	Ile	Lys	Ala	Ala	Ser	Ser	Thr	Glu
		275					280					285			
Lys	Phe	Pro	Asp	Gly	Phe	Trp	Leu	Gly	Glu	Gln	Leu	Val	Cys	Trp	Gln
		290				295					300				

Ala Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu  
305 310 315 320

Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln  
325 330 335

Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys  
340 345 350

Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala  
355 360 365

Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg  
370 375 380

Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr  
385 390 395 400

Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly  
405 410 415

Tyr Asn Ile Pro Gln Thr Asp Glu Ser  
420 425

<210> 29

<211> 1362

<212> DNA

<213> Homo sapiens

<400> 29

```

atggcccaag ccctgccctg gtcctctgctg tggatgggag cgggagtgct gcctgcccac 60
ggcaccacgc acggcatccg gctgccccctg cgcagcggcc tgggggggag cccctggggg 120
ctgaggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
gtgggcagcc ccccgagac gctcaacatc ctggtggata caggcagcag taactttgca 300
gtgggtgctg ccccccaccc ctctctgcat cgctactacc agaggcagct gtccagcaca 360
taccgggacc tccggaaggg tgtgtatgtg ccctacaccc agggcaagtg ggaaggggag 420
ctgggcaccg acctggtgtaag catcccccat ggccccacg tcaactgtgcg tgccaacatt 480
gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcacacctg 540
gggctggcct atgctgagat tgccaggcct gacgactccc tggagccttt ctttgactct 600
ctggtaaagc agaccacagt tcccaacctc ttctccctgc acctttgtgg tgctggcttc 660
ccctcaacc agtctgaagt gctggcctct gtcggaggga gcatgatcat tggaggatc 720
gaccactcgc tgtacacagg cagtctctgg tatacaccca tccggcggga gtggtattat 780
gaggtcatca ttgtgagggt ggagatcaat ggacaggatc tgaaaatgga ctgcaaggag 840
tacaactatg acaagagcat tgtggacagt ggcaccacca accttcgttt gcccaagaaa 900
gtgtttgaag ctgcagtcaa atccatcaag gcagcctcct ccacggagaa gttccctgat 960
ggtttctggc taggagagca gctggtgtgc tggcaagcag gcaccacccc ttggaacatt 1020
ttcccagtca tctcactcta cctaattgggt gaggttacca accagtcctt ccgcatcacc 1080
atccttccgc agcaatacct gcggccagtg gaagatgtgg ccacgtccca agacgactgt 1140
tacaagtttg ccattctaca gtcattccac ggcaactgta tgggagctgt tatcatggag 1200
ggcttctacg ttgtctttga tcgggcccga aaacgaattg gctttgctgt cagcgcttgc 1260
catgtgcacg atgagttcag gacggcagcg gtggaaggcc cttttgtcac cttggacatg 1320
gaagactgtg gctacaacat tccacagaca gatgagtcac ga 1362

```

<210> 30

<211> 453

<212> PRT

<213> Homo sapiens

<400> 30

Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val  
 1 5 10 15  
 Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser  
 20 25 30  
 Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp  
 35 40 45  
 Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val  
 50 55 60  
 Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr  
 65 70 75 80  
 Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser  
 85 90 95  
 Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr  
 100 105 110  
 Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val  
 115 120 125  
 Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp  
 130 135 140  
 Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile  
 145 150 155 160  
 Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp  
 165 170 175  
 Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp  
 180 185 190  
 Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro  
 195 200 205  
 Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln  
 210 215 220  
 Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile  
 225 230 235 240  
 Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg  
 245 250 255  
 Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln  
 260 265 270  
 Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val  
 275 280 285  
 Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala  
 290 295 300  
 Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp  
 305 310 315 320  
 Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr

```
<210> 31
<211> 1380
<212> DNA
<213> Homo sapiens
```

```
<210> 32
<211> 459
<212> PRT
<213> Homo sapiens
```

&lt;400&gt; 32

Met Ala Gln Ala Leu Pro Trp Leu Leu Leu Trp Met Gly Ala Gly Val  
 1 5 10 15  
 Leu Pro Ala His Gly Thr Gln His Gly Ile Arg Leu Pro Leu Arg Ser  
 20 25 30  
 Gly Leu Gly Gly Ala Pro Leu Gly Leu Arg Leu Pro Arg Glu Thr Asp  
 35 40 45  
 Glu Glu Pro Glu Glu Pro Gly Arg Arg Gly Ser Phe Val Glu Met Val  
 50 55 60  
 Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val Glu Met Thr  
 65 70 75 80  
 Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp Thr Gly Ser  
 85 90 95  
 Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu His Arg Tyr  
 100 105 110  
 Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg Lys Gly Val  
 115 120 125  
 Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu Gly Thr Asp  
 130 135 140  
 Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg Ala Asn Ile  
 145 150 155 160  
 Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly Ser Asn Trp  
 165 170 175  
 Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg Pro Asp Asp  
 180 185 190  
 Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr His Val Pro  
 195 200 205  
 Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro Leu Asn Gln  
 210 215 220  
 Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile Gly Gly Ile  
 225 230 235 240  
 Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro Ile Arg Arg  
 245 250 255  
 Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile Asn Gly Gln  
 260 265 270  
 Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys Ser Ile Val  
 275 280 285  
 Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val Phe Glu Ala  
 290 295 300  
 Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys Phe Pro Asp  
 305 310 315 320

Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala Gly Thr Thr  
 325 330 335

Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met Gly Glu Val  
 340 345 350

Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln Tyr Leu Arg  
 355 360 365

Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr Lys Phe Ala  
 370 375 380

Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val Ile Met Glu  
 385 390 395 400

Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile Gly Phe Ala  
 405 410 415

Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala Ala Val Glu  
 420 425 430

Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr Asn Ile Pro  
 435 440 445

Gln Thr Asp Glu Ser His His His His His His  
 450 455

<210> 33

<211> 25

<212> PRT

<213> Homo sapiens

<400> 33

Ser Glu Gln Gln Arg Arg Pro Arg Asp Pro Glu Val Val Asn Asp Glu  
 1 5 10 15

Ser Ser Leu Val Arg His Arg Trp Lys  
 20 25

<210> 34

<211> 19

<212> PRT

<213> Homo sapiens

<400> 34

Ser Glu Gln Leu Arg Gln Gln His Asp Asp Phe Ala Asp Asp Ile Ser  
 1 5 10 15

Leu Leu Lys

<210> 35

<211> 29

<212> DNA

<213> Homo sapiens

<400> 35

gtgatccac ccagcacggc atccggctg

<210> 36  
 <211> 36  
 <212> DNA  
 <213> Homo sapiens

<400> 36  
 gaaagctttc atgactcatc tgtctgtgga atgttg 36

<210> 37  
 <211> 39  
 <212> DNA  
 <213> Homo sapiens

<400> 37  
 gatcgatgac tatctctgac tctccgctg aacaggacg 39

<210> 38  
 <211> 39  
 <212> DNA  
 <213> Homo sapiens

<400> 38  
 gatccgtcct gttcacgcgg agagtcagag atagtcac 39

<210> 39  
 <211> 77  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Hu-Asp2

<400> 39  
 cggcatccgg ctgcccctgc gtagcggctt ggggtggctg cactggggtc tgcgtctgcc 60  
 ccgggagacc gacgaag 77

<210> 40  
 <211> 77  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Hu-Asp2

<400> 40  
 cttcgtcggg ctcccggggc agacgcagac ccagtggagc accacccaga ccgctacgca 60  
 ggggcagccg gatgccg 77

<210> 41  
 <211> 51  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Caspase 8  
 Cleavage Site

<400> 41  
 gatcgatgac tatctctgac tctccgctgg actctggtat cgaaaccgac g 51

<210> 42

<211> 51  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: Caspase 8  
 Cleavage Site  
  
 <400> 42  
 gatccgtcgg tttcgatacc agagtccagc ggagagtcag agatagtcac c 51  
  
 <210> 43  
 <211> 32  
 <212> DNA  
 <213> Homo sapiens  
  
 <400> 43  
 aaggatcctt tgtggagatg gtggacaacc tg 32  
  
 <210> 44  
 <211> 36  
 <212> DNA  
 <213> Homo sapiens  
  
 <400> 44  
 gaaagctttc atgactcatc tgtctgtgga atgttg 36  
  
 <210> 45  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: 6-His tag  
  
 <400> 45  
 gatcgcatca tcaccatcac catg 24  
  
 <210> 46  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: 6-His tag  
  
 <400> 46  
 gatccatggt gatggtgatg atgc 24  
  
 <210> 47  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Description of Artificial Sequence: primer  
  
 <400> 47  
 gactgaccac tcgaccaggt tc 22  
  
 <210> 48  
 <211> 51



<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 48

cgaattaaat tccagcacac tggctacttc ttgttctgca tctcaaagaa c 51

<210> 49

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 49

cgaattaaat tccagcacac tggcta 26

<210> 50

<211> 1287

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Hu-Asp2(b)  
delta TM

<400> 50

```

atggcccaag ccctgccctg gctcctgctg tggatgggag cgggagtgct gcctgcccac 60
ggcaccacag acggcatccg gctgcccctg cgcagcggcc tggggggcgc cccctggggg 120
ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
gtgggcagcc cccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300
gtgggtgctg ccccccaccc ctctctgcat cgctactacc agaggcagct gtccagcaca 360
taccgggacc tccggaaggg tgtgtatgtg cctacacccc agggcaagtg ggaagggggag 420
ctgggcaccg acctggtaag catcccccat ggccccaacg tcaactgtgcg tgccaacatt 480
gctgccatca ctgaatcaga caagttcttc atcaacggct ccaactggga aggcacacct 540
gggctggcct atgtgagat tgccaggctt tgtgtgctg gcttccccct caaccagtct 600
gaagtgtgag cctctgtcgg agggagcatg atcattggag gtatcgacca ctcgctgtac 660
acaggcagtc tctgtgtatc acccatccgg cgggagtggt attatgaggt catcattgtg 720
cgggtggaga tcaatggaca ggatctgaaa atggactgca aggagtacaa ctatgacaag 780
agcattgtgg acagtggcac caccaacctt cgtttgccca agaaagtgtt tgaagctgca 840
gtcaaatacca tcaaggcagc ctctccacg gagaagttcc ctgatggttt ctggctagga 900
gagcagctgg tgtgctggca agcaggcacc accccttgga acattttccc agtcacatca 960
ctctacctaa tgggtgaggt taccaaccag tccttcgcga tcaccatcct tccgcagcaa 1020
tacctgcggc cagtgggaaga tgtggccacg tcccaagacg actgttacaa gtttgccatc 1080
tcacagtcat ccacgggcac tgttatggga gctgttatca tggaggggctt ctacgttgct 1140
tttgatcggg ccggaacacg aattggcttt gctgtcagcg cttgccatgt gcacgatgag 1200
ttcaggacgg cagcgggtgga aggccctttt gtcaccttgg acatggaaga ctgtggctac 1260
aacattccac agacagatga gtcatga 1287

```

<210> 51

<211> 428

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Hu-Asp2(b)  
delta TM

&lt;400&gt; 51

Met	Ala	Gln	Ala	Leu	Pro	Trp	Leu	Leu	Leu	Trp	Met	Gly	Ala	Gly	Val
1				5					10					15	
Leu	Pro	Ala	His	Gly	Thr	Gln	His	Gly	Ile	Arg	Leu	Pro	Leu	Arg	Ser
			20					25					30		
Gly	Leu	Gly	Gly	Ala	Pro	Leu	Gly	Leu	Arg	Leu	Pro	Arg	Glu	Thr	Asp
	35						40					45			
Glu	Glu	Pro	Glu	Glu	Pro	Gly	Arg	Arg	Gly	Ser	Phe	Val	Glu	Met	Val
	50					55					60				
Asp	Asn	Leu	Arg	Gly	Lys	Ser	Gly	Gln	Gly	Tyr	Tyr	Val	Glu	Met	Thr
65					70					75					80
Val	Gly	Ser	Pro	Pro	Gln	Thr	Leu	Asn	Ile	Leu	Val	Asp	Thr	Gly	Ser
				85					90					95	
Ser	Asn	Phe	Ala	Val	Gly	Ala	Ala	Pro	His	Pro	Phe	Leu	His	Arg	Tyr
			100					105					110		
Tyr	Gln	Arg	Gln	Leu	Ser	Ser	Thr	Tyr	Arg	Asp	Leu	Arg	Lys	Gly	Val
	115						120					125			
Tyr	Val	Pro	Tyr	Thr	Gln	Gly	Lys	Trp	Glu	Gly	Glu	Leu	Gly	Thr	Asp
	130					135					140				
Leu	Val	Ser	Ile	Pro	His	Gly	Pro	Asn	Val	Thr	Val	Arg	Ala	Asn	Ile
145					150					155					160
Ala	Ala	Ile	Thr	Glu	Ser	Asp	Lys	Phe	Phe	Ile	Asn	Gly	Ser	Asn	Trp
				165					170					175	
Glu	Gly	Ile	Leu	Gly	Leu	Ala	Tyr	Ala	Glu	Ile	Ala	Arg	Leu	Cys	Gly
			180					185					190		
Ala	Gly	Phe	Pro	Leu	Asn	Gln	Ser	Glu	Val	Leu	Ala	Ser	Val	Gly	Gly
	195						200					205			
Ser	Met	Ile	Ile	Gly	Gly	Ile	Asp	His	Ser	Leu	Tyr	Thr	Gly	Ser	Leu
	210					215					220				
Trp	Tyr	Thr	Pro	Ile	Arg	Arg	Glu	Trp	Tyr	Tyr	Glu	Val	Ile	Ile	Val
225					230					235					240
Arg	Val	Glu	Ile	Asn	Gly	Gln	Asp	Leu	Lys	Met	Asp	Cys	Lys	Glu	Tyr
				245					250					255	
Asn	Tyr	Asp	Lys	Ser	Ile	Val	Asp	Ser	Gly	Thr	Thr	Asn	Leu	Arg	Leu
		260						265					270		
Pro	Lys	Lys	Val	Phe	Glu	Ala	Ala	Val	Lys	Ser	Ile	Lys	Ala	Ala	Ser
		275					280					285			
Ser	Thr	Glu	Lys	Phe	Pro	Asp	Gly	Phe	Trp	Leu	Gly	Glu	Gln	Leu	Val
	290					295					300				
Cys	Trp	Gln	Ala	Gly	Thr	Thr	Pro	Trp	Asn	Ile	Phe	Pro	Val	Ile	Ser
305					310					315					320

Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile  
 325 330 335  
 Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln  
 340 345 350  
 Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val  
 355 360 365  
 Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala  
 370 375 380  
 Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu  
 385 390 395 400  
 Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu  
 405 410 415  
 Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser  
 420 425

<210> 52

<211> 1305

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Hu-Asp2(b)  
delta TM

<400> 52

```

atggcccaag ccctgccctg gctcctgctg tggatgggag cgggagtgtc gcctgcccac 60
ggcaccagc acggcatccg gctgcccctg cgcagcggcc tggggggcgc cccctgggg 120
ctgcggctgc cccgggagac cgacgaagag cccgaggagc ccggccggag gggcagcttt 180
gtggagatgg tggacaacct gaggggcaag tcggggcagg gctactacgt ggagatgacc 240
gtgggcagcc cccgcagac gctcaacatc ctggtggata caggcagcag taactttgca 300
gtgggtgctg ccccccaccc cttcctgcat cgctactacc agaggcagct gtccagcaca 360
taccgggacc tccggaaggg tgtgtatgtg cctacaccc agggcaagtg ggaaggggag 420
ctgggcaccg acctggtaag catcccccat ggccccaacg tcaactgtgc tgccaacatt 480
gctgccatca ctgaatcaga caagtctctc atcaacggct ccaactggga aggcattctg 540
gggctggcct atgctgagat tgccaggctt tgtgtgtgtg gcttccccct caaccagtct 600
gaagtgtgtg cctctgtcgg agggagcatg atcattggag gtatcgacca ctgctgttac 660
acaggcagtc tctgtgtatc acccatccgg cgggagtggg attatgaggt catcattgtg 720
cgggtggaga tcaatggaca ggaatctgaaa atggactgca aggagtacaa ctatgacaag 780
agcattgtgg acagtggcac caccaacctt cgtttgccca agaaagtgtt tgaagctgca 840
gtcaaatcca tcaaggcagc ctccctccacg gagaagttcc ctgatggttt ctggctagga 900
gagcagctgg tgtgctggca agcaggcacc accccttgga acattttccc agtcatctca 960
ctctacctaa tgggtgaggt taccaaccag tccttcgcga tcaccatcct tccgcagcaa 1020
tacctgctgc cagtgaaga tgtggccacg tcccaagacg actgttataa gtttgccatc 1080
tcacagtcac ccacgggcac tggtatggga gctgttatca tggagggtt ctacgttgct 1140
tttgatcggg ccgaaaaacg aattggcttt gctgtcagcg cttgccatgt gcacgatgag 1200
ttcaggacgg cagcgggtga aggccctttt gtcaccttgg acatggaaga ctgtggctac 1260
aacattccac agacagatga gtcacagcag cagcagcagc agtga 1305

```

<210> 53

<211> 434

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Hu-Asp2(b)

## delta TM

&lt;400&gt; 53

Met	Ala	Gln	Ala	Leu	Pro	Trp	Leu	Leu	Leu	Trp	Met	Gly	Ala	Gly	Val	1	5	10	15
Leu	Pro	Ala	His	Gly	Thr	Gln	His	Gly	Ile	Arg	Leu	Pro	Leu	Arg	Ser	20	25	30	
Gly	Leu	Gly	Gly	Ala	Pro	Leu	Gly	Leu	Arg	Leu	Pro	Arg	Glu	Thr	Asp	35	40	45	
Glu	Glu	Pro	Glu	Glu	Pro	Gly	Arg	Arg	Gly	Ser	Phe	Val	Glu	Met	Val	50	55	60	
Asp	Asn	Leu	Arg	Gly	Lys	Ser	Gly	Gln	Gly	Tyr	Tyr	Val	Glu	Met	Thr	65	70	75	80
Val	Gly	Ser	Pro	Pro	Gln	Thr	Leu	Asn	Ile	Leu	Val	Asp	Thr	Gly	Ser	85	90	95	
Ser	Asn	Phe	Ala	Val	Gly	Ala	Ala	Pro	His	Pro	Phe	Leu	His	Arg	Tyr	100	105	110	
Tyr	Gln	Arg	Gln	Leu	Ser	Ser	Thr	Tyr	Arg	Asp	Leu	Arg	Lys	Gly	Val	115	120	125	
Tyr	Val	Pro	Tyr	Thr	Gln	Gly	Lys	Trp	Glu	Gly	Glu	Leu	Gly	Thr	Asp	130	135	140	
Leu	Val	Ser	Ile	Pro	His	Gly	Pro	Asn	Val	Thr	Val	Arg	Ala	Asn	Ile	145	150	155	160
Ala	Ala	Ile	Thr	Glu	Ser	Asp	Lys	Phe	Phe	Ile	Asn	Gly	Ser	Asn	Trp	165	170	175	
Glu	Gly	Ile	Leu	Gly	Leu	Ala	Tyr	Ala	Glu	Ile	Ala	Arg	Leu	Cys	Gly	180	185	190	
Ala	Gly	Phe	Pro	Leu	Asn	Gln	Ser	Glu	Val	Leu	Ala	Ser	Val	Gly	Gly	195	200	205	
Ser	Met	Ile	Ile	Gly	Gly	Ile	Asp	His	Ser	Leu	Tyr	Thr	Gly	Ser	Leu	210	215	220	
Trp	Tyr	Thr	Pro	Ile	Arg	Arg	Glu	Trp	Tyr	Tyr	Glu	Val	Ile	Ile	Val	225	230	235	240
Arg	Val	Glu	Ile	Asn	Gly	Gln	Asp	Leu	Lys	Met	Asp	Cys	Lys	Glu	Tyr	245	250	255	
Asn	Tyr	Asp	Lys	Ser	Ile	Val	Asp	Ser	Gly	Thr	Thr	Asn	Leu	Arg	Leu	260	265	270	
Pro	Lys	Lys	Val	Phe	Glu	Ala	Ala	Val	Lys	Ser	Ile	Lys	Ala	Ala	Ser	275	280	285	
Ser	Thr	Glu	Lys	Phe	Pro	Asp	Gly	Phe	Trp	Leu	Gly	Glu	Gln	Leu	Val	290	295	300	
Cys	Trp	Gln	Ala	Gly	Thr	Thr	Pro	Trp	Asn	Ile	Phe	Pro	Val	Ile	Ser	305	310	315	320

Leu Tyr Leu Met Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile  
325 330 335

Leu Pro Gln Gln Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln  
340 345 350

Asp Asp Cys Tyr Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val  
355 360 365

Met Gly Ala Val Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala  
370 375 380

Arg Lys Arg Ile Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu  
385 390 395 400

Phe Arg Thr Ala Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu  
405 410 415

Asp Cys Gly Tyr Asn Ile Pro Gln Thr Asp Glu Ser His His His His  
420 425 430

His His

<210> 54

<211> 2310

<212> DNA

<213> Homo sapiens

<400> 54

```

atgctgcccc gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggta 60
ccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaaa 180
acctgcattg ataccaagga aggcattcctg cagtattgcc aagaagtcta ccctgaactg 240
cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgaagac ccattcccac tttgtgatc cctaccgctg cttagtgtgt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgc aattcttaca ccaggagagg 420
atggatgttt gcgaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagttag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagt acaatgtgga ttctgctgat 600
gcgaggagg atgactcggg tgtctgtgtg ggaggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgcca ccaccaccac caccaccaca 840
gagtcctgtg aagagtggtg tcgagaggtg tgctctgaac aagccgagac ggggccgtgc 900
cgagcaatga tctcccgctg gtactttgat gtgactgaag ggaagtgtgc ccattcttt 960
tacggcggat gtggcggcaa ccggaacaac tttgacacag aagagtactg catggccgtg 1020
tgtggcagcg ccatgtccca aagtttactc aagactaccc aggaacctct tggccgagat 1080
cctgttaaac ttcctacaac agcagccagt acccctgatg ccgttgacaa gtatctcgag 1140
acacctgggg atgagaatga acatgcccat ttccagaaag ccaaagagag gcttgaggcc 1200
aagcaccgag agagaatgtc ccaggtcatg agagaatggg aagaggcaga acgtcaagca 1260
aagaacttgc cttaaagctg taagaaggca gttatccagc atttccagga gaaagtgaa 1320
tctttggaac aggaagcagc caacgagaga cagcagctgg tggagacaca catggccaga 1380
gtggaagcca tgetcaatga ccgccgccc ctggccctgg agaactacat caccgctctg 1440
caggctgttc ctctcggccc tcgtcacgtg ttcaatatgc taaagaagta tgtccgcgca 1500
gaacagaagg acagacagca caccctaaag catttcgagc atgtgcgcat ggtggatccc 1560
aagaaagccg ctcatgccg gtcccagggt atgacacacc tccgtgtgat ttatgagcgc 1620
atgaatcagt ctctctccct gctctacaac gtgcctgcag tggccgagga gattcaggat 1680
gaagttgatg agctgcttca gaaagagcaa aactattcag atgacgtctt ggccaacatg 1740

```

```

attagtgaac caaggatcag ttacggaaac gatgctctca tgccatcttt gaccgaaacg 1800
aaaaccaccg tggagctcct tcccgtgaat ggagagttca gcctggacga tctccagccg 1860
tggcattctt ttggggctga ctctgtgcca gccaacacag aaaacgaagt tgagcctgtt 1920
gatgcccgcc ctgctgccga ccgaggactg accactcgac caggttctgg gttgacaaat 1980
atcaagacgg aggagatctc tgaagtgaag atggatgcag aattccgaca tgactcagga 2040
tatgaagttc atcatcaaaa attggtgttc tttgcagaag atgtgggttc aaacaaaggt 2100
gcaatcattg gactcatggt gggcgggtgtt gtcatagcga cagtgatcgt catcaccttg 2160
gtgatgctga agaagaaaca gtacacatcc attcatcatg gtgtggtgga gggtgacgcc 2220
gctgtcaccg cagaggagcg ccacctgtcc aagatgcagc agaacggcta cgaaaatcca 2280
acctacaagt tctttgagca gatgcagaac 2310

```

&lt;210&gt; 55

&lt;211&gt; 770

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 55

```

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
  1              5              10              15

```

```

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
      20              25              30

```

```

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
      35              40              45

```

```

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
      50              55              60

```

```

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
      65              70              75              80

```

```

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
      85              90              95

```

```

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
      100              105              110

```

```

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
      115              120              125

```

```

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
      130              135              140

```

```

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
      145              150              155              160

```

```

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile
      165              170              175

```

```

Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu
      180              185              190

```

```

Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val
      195              200              205

```

```

Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys
      210              215              220

```

```

Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu
      225              230              235              240

```

Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu  
 245 250 255  
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile  
 260 265 270  
 Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg  
 275 280 285  
 Glu Val Cys Ser Glu Gln Ala Glu Thr Gly Pro Cys Arg Ala Met Ile  
 290 295 300  
 Ser Arg Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys Ala Pro Phe Phe  
 305 310 315 320  
 Tyr Gly Gly Cys Gly Gly Asn Arg Asn Asn Phe Asp Thr Glu Glu Tyr  
 325 330 335  
 Cys Met Ala Val Cys Gly Ser Ala Met Ser Gln Ser Leu Leu Lys Thr  
 340 345 350  
 Thr Gln Glu Pro Leu Ala Arg Asp Pro Val Lys Leu Pro Thr Thr Ala  
 355 360 365  
 Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu Glu Thr Pro Gly Asp  
 370 375 380  
 Glu Asn Glu His Ala His Phe Gln Lys Ala Lys Glu Arg Leu Glu Ala  
 385 390 395 400  
 Lys His Arg Glu Arg Met Ser Gln Val Met Arg Glu Trp Glu Glu Ala  
 405 410 415  
 Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp Lys Lys Ala Val Ile  
 420 425 430  
 Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu Gln Glu Ala Ala Asn  
 435 440 445  
 Glu Arg Gln Gln Leu Val Glu Thr His Met Ala Arg Val Glu Ala Met  
 450 455 460  
 Leu Asn Asp Arg Arg Arg Leu Ala Leu Glu Asn Tyr Ile Thr Ala Leu  
 465 470 475 480  
 Gln Ala Val Pro Pro Arg Pro Arg His Val Phe Asn Met Leu Lys Lys  
 485 490 495  
 Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His Thr Leu Lys His Phe  
 500 505 510  
 Glu His Val Arg Met Val Asp Pro Lys Lys Ala Ala Gln Ile Arg Ser  
 515 520 525  
 Gln Val Met Thr His Leu Arg Val Ile Tyr Glu Arg Met Asn Gln Ser  
 530 535 540  
 Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala Glu Glu Ile Gln Asp  
 545 550 555 560  
 Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn Tyr Ser Asp Asp Val

<400> 56						
atgctgcccg	gtttggcact	gctcctgctg	gccgcctgga	cggctcgggc	gctggaggta	60
cccactgatg	gtaatgctgg	cctgctggct	gaaccccaga	tgccatgtt	ctgtggcaga	120
ctgaacatgc	acatgaatgt	ccagaatggg	aagtgggatt	cagatccatc	agggaccaa	180
acctgcattg	ataccaagga	aggcaatcctg	cagtattgcc	aagaagtcta	ccctgaactg	240
cagataccga	atgtggtaga	agccaaccaa	ccagtgacca	tccagaactg	gtgcaagcgg	300
ggccgcaagc	agtgcaagac	ccatccccac	tttgtgattc	cctaccgctg	cttaagtggg	360
gagtttgtaa	gtgatgccct	tctcgttcct	gacaagtgca	aattcttaca	ccaggagagg	420
atggatgttt	gcgaaactca	tcttcaactgg	cacaccgtcg	ccaaagagac	atgcagtgag	480
aagagtacca	acttgcgatg	ctacggcgatg	ttgctgccct	gcggaattga	caagttccga	540
ggggtagagt	ttgtgtgttg	cccactggct	gaagaaagtg	acaatgtgga	ttctgctgat	600
gctggaggagg	atgactcgga	tgtctggttg	ggcggagcag	acacagacta	tgcatagggg	660
agtgaagaca	aagttagtaga	agttagcagag	gaggagaaga	ttgctgaggt	ggaagaagaa	720
gaagccgatg	atgacgagga	cgatgaggat	ggtgatgagg	taggaggaaga	ggctgaggaa	780
ccctacgaag	aagccacaga	gagaaccacc	agcattgccca	ccaccaccac	caccaccaca	840



```

gagtctgtgg aagaggtggt tgcagaggtg tgctctgaac aagccgagac ggggccgtgc 900
cgagcaatga tctcccgtg gtactttgat gtgactgaag ggaagtgtgc cccattcttt 960
tacggcggat gtggcggcaa ccggaacaac tttgacacag aagagtactg catggccgtg 1020
tgtggcagcg ccattcctac aacagcagcc agtaccctg atgccgttga caagtatctc 1080
gagacacctg gggatgagaa tgaacatgcc catttccaga aagccaaaga gaggcttgag 1140
gccaagcacc gagagagaat gtcccaggtc atgagagaat ggaagaggc agaacgtcaa 1200
gcaaagaact tgcctaaagc tgataagaag gcagttatcc agcatttcca ggagaaagtg 1260
gaatcttttg aacaggaagc agccaacgag agacagcagc tgggtggagac acacatggcc 1320
agagtgggaag ccatgctcaa tgaccgccgc cgctgggcc tggagaacta catcacgct 1380
ctgcaggctg ttctcctcgc gctcgtcac gtgttcaata tgctaaagaa gtatgtccgc 1440
gcagaacaga aggacagaca gcacacccta aagcatttcg agcatgtgcg catggtggat 1500
cccaagaaag ccgctcagat ccggtcccag gttatgacac acctccgtgt gatttatgag 1560
cgcatgaatc agtctctctc cctgctctac aacgtgcctg cagtggccga ggagattcag 1620
gatgaagttg atgagctgct tcagaaagag caaaactatt cagatgacgt cttggccaac 1680
atgattagtg aaccaaggat cagttacgga aacgatgctc tcatgccatc tttgaccgaa 1740
acgaaaacca ccgtggagct ccttcccgtg aatggagagt tcagcctgga cgatctccag 1800
ccgtggcatt cttttggggc tgactctgtg ccagccaaca cagaaaacga agttgagcct 1860
gttgatgccc gccctgctgc cgaccgagga ctgaccactc gaccagggtc tgggttgaca 1920
aatatcaaga cggaggagat ctctgaagtg aagatggatg cagaattccg acatgactca 1980
ggatatgaag ttcacatca aaaattggtg ttctttgcag aagatgtggg ttcaaacaaa 2040
ggtgcaatca ttgactcat ggtggggcgt gttgtcatag cgacagtgt cgtcatcacc 2100
ttggtgatgc tgaagaagaa acagtacaca tccattcatc atggtgtggt ggaggttgac 2160
gccgctgtca cccagagga gcgccacctg tccaagatgc agcagaacgg ctacgaaaat 2220
ccaacctaca agttctttga gcagatgcag aac 2253

```

<210> 57

<211> 751

<212> PRT

<213> Homo sapiens

<400> 57

```

Met Leu Pro Gly Leu Ala Leu Leu Leu Leu Ala Ala Trp Thr Ala Arg
 1              5              10              15

```

```

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
      20              25              30

```

```

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
      35              40              45

```

```

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
      50              55              60

```

```

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
      65              70              75              80

```

```

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
      85              90              95

```

```

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
      100              105              110

```

```

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu
      115              120              125

```

```

Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys
      130              135              140

```

```

Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu
      145              150              155              160

```

Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile  
 165 170 175  
 Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu  
 180 185 190  
 Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val  
 195 200 205  
 Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys  
 210 215 220  
 Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu  
 225 230 235 240  
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu  
 245 250 255  
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile  
 260 265 270  
 Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg  
 275 280 285  
 Glu Val Cys Ser Glu Gln Ala Glu Thr Gly Pro Cys Arg Ala Met Ile  
 290 295 300  
 Ser Arg Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys Ala Pro Phe Phe  
 305 310 315 320  
 Tyr Gly Gly Cys Gly Gly Asn Arg Asn Asn Phe Asp Thr Glu Glu Tyr  
 325 330 335  
 Cys Met Ala Val Cys Gly Ser Ala Ile Pro Thr Thr Ala Ala Ser Thr  
 340 345 350  
 Pro Asp Ala Val Asp Lys Tyr Leu Glu Thr Pro Gly Asp Glu Asn Glu  
 355 360 365  
 His Ala His Phe Gln Lys Ala Lys Glu Arg Leu Glu Ala Lys His Arg  
 370 375 380  
 Glu Arg Met Ser Gln Val Met Arg Glu Trp Glu Glu Ala Glu Arg Gln  
 385 390 395 400  
 Ala Lys Asn Leu Pro Lys Ala Asp Lys Lys Ala Val Ile Gln His Phe  
 405 410 415  
 Gln Glu Lys Val Glu Ser Leu Glu Gln Glu Ala Ala Asn Glu Arg Gln  
 420 425 430  
 Gln Leu Val Glu Thr His Met Ala Arg Val Glu Ala Met Leu Asn Asp  
 435 440 445  
 Arg Arg Arg Leu Ala Leu Glu Asn Tyr Ile Thr Ala Leu Gln Ala Val  
 450 455 460  
 Pro Pro Arg Pro Arg His Val Phe Asn Met Leu Lys Lys Tyr Val Arg  
 465 470 475 480  
 Ala Glu Gln Lys Asp Arg Gln His Thr Leu Lys His Phe Glu His Val

485					490					495					
Arg	Met	Val	Asp	Pro	Lys	Lys	Ala	Ala	Gln	Ile	Arg	Ser	Gln	Val	Met
			500					505					510		
Thr	His	Leu	Arg	Val	Ile	Tyr	Glu	Arg	Met	Asn	Gln	Ser	Leu	Ser	Leu
		515					520					525			
Leu	Tyr	Asn	Val	Pro	Ala	Val	Ala	Glu	Glu	Ile	Gln	Asp	Glu	Val	Asp
		530					535					540			
Glu	Leu	Leu	Gln	Lys	Glu	Gln	Asn	Tyr	Ser	Asp	Asp	Val	Leu	Ala	Asn
		545					550					555			560
Met	Ile	Ser	Glu	Pro	Arg	Ile	Ser	Tyr	Gly	Asn	Asp	Ala	Leu	Met	Pro
				565					570						575
Ser	Leu	Thr	Glu	Thr	Lys	Thr	Thr	Val	Glu	Leu	Leu	Pro	Val	Asn	Gly
			580					585						590	
Glu	Phe	Ser	Leu	Asp	Asp	Leu	Gln	Pro	Trp	His	Ser	Phe	Gly	Ala	Asp
		595					600					605			
Ser	Val	Pro	Ala	Asn	Thr	Glu	Asn	Glu	Val	Glu	Pro	Val	Asp	Ala	Arg
		610					615					620			
Pro	Ala	Ala	Asp	Arg	Gly	Leu	Thr	Thr	Arg	Pro	Gly	Ser	Gly	Leu	Thr
							630					635			640
Asn	Ile	Lys	Thr	Glu	Glu	Ile	Ser	Glu	Val	Lys	Met	Asp	Ala	Glu	Phe
				645					650					655	
Arg	His	Asp	Ser	Gly	Tyr	Glu	Val	His	His	Gln	Lys	Leu	Val	Phe	Phe
			660					665						670	
Ala	Glu	Asp	Val	Gly	Ser	Asn	Lys	Gly	Ala	Ile	Ile	Gly	Leu	Met	Val
		675					680					685			
Gly	Gly	Val	Val	Ile	Ala	Thr	Val	Ile	Val	Ile	Thr	Leu	Val	Met	Leu
		690					695					700			
Lys	Lys	Lys	Gln	Tyr	Thr	Ser	Ile	His	His	Gly	Val	Val	Glu	Val	Asp
		705					710					715			720
Ala	Ala	Val	Thr	Pro	Glu	Glu	Arg	His	Leu	Ser	Lys	Met	Gln	Gln	Asn
				725					730						735
Gly	Tyr	Glu	Asn	Pro	Thr	Tyr	Lys	Phe	Phe	Glu	Gln	Met	Gln	Asn	
			740					745					750		

&lt;210&gt; 58

&lt;211&gt; 2316

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 58

```

atgctgcccg gtttggcact gtcctgctg gccgcctgga cggctcgggc gctggaggta 60
cccactgatg gtaatgctgg cctgctggct gaacccaga ttgcatgtt ctgtggcaga 120
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggacaaa 180
acctgcattg ataccaagga aggcacctg cagtattgcc aagaagtcta ccctgaactg 240

```

```

cagatcacca atgtggtaga agccaaccaa ccagtgacca tccagaactg gtgcaagcgg 300
ggccgcaagc agtgcaagac ccatccccac tttgtgattc cctaccgctg cttagttagt 360
gagtttgtaa gtgatgccct tctcgttcct gacaagtgca aattcttaca ccaggagagg 420
atggatgttt gcgaaaactca tcttcactgg cacaccgtcg ccaaagagac atgcagtggag 480
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600
gcggaggagg atgactcgga tgtctggtgg ggccggagcag acacagacta tgcagatggg 660
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780
ccctacgaag aagccacaga gagaaccacc agcattgccca ccaccaccac caccaccaca 840
gagtctgtgg aagaggtggt tcgagagggtg tgctctgaac aagccgagac ggggccgtgc 900
cgagcaatga tctcccgctg gtactttgat gtgactgaag ggaagtgtgc cccattcttt 960
tacggcggat gtggcgccaa ccggaacaac tttgacacag aagagtactg catggccgtg 1020
tgtggcagcg ccatgtccca aagtttactc aagactaccc aggaacctct tggccgagat 1080
cctgttaaac ttcctacaac agcagccagt acccctgatg ccgttgacaa gtatctcgag 1140
acacctgggg atgagaatga acatgcccac ttccagaaag ccaaagagag gcttgaggcc 1200
aagcaccgag agagaatgtc ccaggtcatg agagaatggg aagaggcaga acgtcaagca 1260
aagaacttgc ctaaagctga taagaaggca gttatccagc atttccagga gaaagtggaa 1320
tctttggaac aggaagcagc caacgagaga cagcagctgg tggagacaca catggccaga 1380
gtggaagcca tgctcaatga ccgccgccgc ctggccctgg agaactacat caccgctctg 1440
caggctgttc ctcctcgcc tcgtcacgtg ttcaatatgc taaagaagta tgtccgcgca 1500
gaacagaagg acagacagca caccctaaag catttcgagc atgtgcgcat ggtggatccc 1560
aagaaagccg ctcagatccg gtcccagggt atgacacacc tccgtgtgat ttatgagcgc 1620
atgaatcagt ctctctccct gctctacaac gtgcctgcag tggccgagga gattcaggat 1680
gaagttgatg agctgcttca gaaagagcaa aactattcag atgacgtctt ggccaacatg 1740
attagtgaac caaggatcag ttacggaaac gatgctctca tgccatcttt gaccgaaacg 1800
aaaaccaccg tggagctcct tcccgtgaat ggagagttca gcctggacga tctccagccg 1860
tggcattctt ttggggctga ctctgtgccca gccaacacag aaaacgaagt tgagcctgtt 1920
gatgcccggc ctgctgccga ccgaggactg accactcgac caggttctgg gttgacaaat 1980
atcaagacgg aggagatctc tgaagtgaag atggatgcag aattccgaca tgactcagga 2040
tatgaagttc atcatcaaaa attggtgttc tttgcagaag atgtgggttc aaacaaaggt 2100
gcaatcattg gactcatggt gggcggtgtt gtcatagcga cagtgatcgt catcaccttg 2160
gtgatgctga agaagaaaca gtacacatcc attcatcatg gtgtggtgga ggttgacgcc 2220
gctgtcaccc cagaggagcg ccacctgtcc aagatgcagc agaacggcta cgaaaatcca 2280
acctacaagt tctttgagca gatgcagaac aagaag 2316

```

&lt;210&gt; 59

&lt;211&gt; 772

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 59

```

Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala Trp Thr Ala Arg
  1             5             10             15

```

```

Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro
  20             25             30

```

```

Gln Ile Ala Met Phe Cys Gly Arg Leu Asn Met His Met Asn Val Gln
  35             40             45

```

```

Asn Gly Lys Trp Asp Ser Asp Pro Ser Gly Thr Lys Thr Cys Ile Asp
  50             55             60

```

```

Thr Lys Glu Gly Ile Leu Gln Tyr Cys Gln Glu Val Tyr Pro Glu Leu
  65             70             75             80

```

```

Gln Ile Thr Asn Val Val Glu Ala Asn Gln Pro Val Thr Ile Gln Asn
  85             90             95

```

```

Trp Cys Lys Arg Gly Arg Lys Gln Cys Lys Thr His Pro His Phe Val
 100            105            110

```

Ile Pro Tyr Arg Cys Leu Val Gly Glu Phe Val Ser Asp Ala Leu Leu  
 115 120 125  
 Val Pro Asp Lys Cys Lys Phe Leu His Gln Glu Arg Met Asp Val Cys  
 130 135 140  
 Glu Thr His Leu His Trp His Thr Val Ala Lys Glu Thr Cys Ser Glu  
 145 150 155 160  
 Lys Ser Thr Asn Leu His Asp Tyr Gly Met Leu Leu Pro Cys Gly Ile  
 165 170 175  
 Asp Lys Phe Arg Gly Val Glu Phe Val Cys Cys Pro Leu Ala Glu Glu  
 180 185 190  
 Ser Asp Asn Val Asp Ser Ala Asp Ala Glu Glu Asp Asp Ser Asp Val  
 195 200 205  
 Trp Trp Gly Gly Ala Asp Thr Asp Tyr Ala Asp Gly Ser Glu Asp Lys  
 210 215 220  
 Val Val Glu Val Ala Glu Glu Glu Glu Val Ala Glu Val Glu Glu Glu  
 225 230 235 240  
 Glu Ala Asp Asp Asp Glu Asp Asp Glu Asp Gly Asp Glu Val Glu Glu  
 245 250 255  
 Glu Ala Glu Glu Pro Tyr Glu Glu Ala Thr Glu Arg Thr Thr Ser Ile  
 260 265 270  
 Ala Thr Thr Thr Thr Thr Thr Thr Glu Ser Val Glu Glu Val Val Arg  
 275 280 285  
 Glu Val Cys Ser Glu Gln Ala Glu Thr Gly Pro Cys Arg Ala Met Ile  
 290 295 300  
 Ser Arg Trp Tyr Phe Asp Val Thr Glu Gly Lys Cys Ala Pro Phe Phe  
 305 310 315 320  
 Tyr Gly Gly Cys Gly Gly Asn Arg Asn Asn Phe Asp Thr Glu Glu Tyr  
 325 330 335  
 Cys Met Ala Val Cys Gly Ser Ala Met Ser Gln Ser Leu Leu Lys Thr  
 340 345 350  
 Thr Gln Glu Pro Leu Ala Arg Asp Pro Val Lys Leu Pro Thr Thr Ala  
 355 360 365  
 Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr Leu Glu Thr Pro Gly Asp  
 370 375 380  
 Glu Asn Glu His Ala His Phe Gln Lys Ala Lys Glu Arg Leu Glu Ala  
 385 390 395 400  
 Lys His Arg Glu Arg Met Ser Gln Val Met Arg Glu Trp Glu Glu Ala  
 405 410 415  
 Glu Arg Gln Ala Lys Asn Leu Pro Lys Ala Asp Lys Lys Ala Val Ile  
 420 425 430  
 Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu Gln Glu Ala Ala Asn

435	440	445
Glu Arg Gln Gln Leu Val	Glu Thr His Met Ala Arg Val	Glu Ala Met
450	455	460
Leu Asn Asp Arg Arg Arg	Leu Ala Leu Glu Asn Tyr Ile Thr Ala Leu	
465	470	475 480
Gln Ala Val Pro Pro Arg Pro Arg His	Val Phe Asn Met Leu Lys Lys	
	485	490 495
Tyr Val Arg Ala Glu Gln Lys Asp Arg	Gln His Thr Leu Lys His Phe	
	500	505 510
Glu His Val Arg Met Val Asp Pro Lys Lys Ala Ala	Gln Ile Arg Ser	
	515	520 525
Gln Val Met Thr His Leu Arg Val Ile Tyr Glu Arg Met Asn Gln Ser		
	530	535 540
Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala Glu Glu Ile Gln Asp		
	545	550 555 560
Glu Val Asp Glu Leu Leu Gln Lys Glu Gln Asn Tyr Ser Asp Asp Val		
	565	570 575
Leu Ala Asn Met Ile Ser Glu Pro Arg Ile Ser Tyr Gly Asn Asp Ala		
	580	585 590
Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr Val Glu Leu Leu Pro		
	595	600 605
Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln Pro Trp His Ser Phe		
	610	615 620
Gly Ala Asp Ser Val Pro Ala Asn Thr Glu Asn Glu Val Glu Pro Val		
	625	630 635 640
Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr Thr Arg Pro Gly Ser		
	645	650 655
Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser Glu Val Lys Met Asp		
	660	665 670
Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys Leu		
	675	680 685
Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly		
	690	695 700
Leu Met Val Gly Gly Val Val Ile Ala Thr Val Ile Val Ile Thr Leu		
	705	710 715 720
Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile His His Gly Val Val		
	725	730 735
Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met		
	740	745 750
Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met		

755

760

765

Gln Asn Lys Lys  
770

<210> 60  
<211> 2259  
<212> DNA  
<213> Homo sapiens

<400> 60  
atgctgcccc gtttggcact gctcctgctg gccgcctgga cggctcgggc gctggaggtta 60  
cccactgatg gtaatgctgg cctgctggct gaaccccaga ttgccatgtt ctgtggcaga 120  
ctgaacatgc acatgaatgt ccagaatggg aagtgggatt cagatccatc agggaccaa 180  
acctgcattg ataccaagga aggcacctg cagtattgcc aagaagtcta ccctgaactg 240  
cagatcacca atgtggtaga agccaacca ccagtgaacca tccagaactg gtgcaagcgg 300  
ggccgcaagc agtgcaagac ccatcccccac tttgtgattc cctaccgctg cttagtgtgg 360  
gagtttgtaa gtgatgccct tctcgttccct gacaagtga aattcttaca ccaggagagg 420  
atggatgttt gcgaaactca tcttccactgg cacaccgtcg ccaaagagac atgcagttag 480  
aagagtacca acttgcatga ctacggcatg ttgctgccct gcggaattga caagttccga 540  
ggggtagagt ttgtgtgttg cccactggct gaagaaagtg acaatgtgga ttctgctgat 600  
gctggaggagg atgactcgga tgtctggtgg ggccggagcag acacagacta tgcagatggg 660  
agtgaagaca aagtagtaga agtagcagag gaggaagaag tggctgaggt ggaagaagaa 720  
gaagccgatg atgacgagga cgatgaggat ggtgatgagg tagaggaaga ggctgaggaa 780  
ccctacgaag aagccacaga gagaaccacc agcattgccca ccaccaccac caccaccaca 840  
gagtcctgtg aagaggtggt tgcgagagtg tgctctgaac aagccgagac ggggccgtgc 900  
cgagcaatga tctcccgctg gtactttgat gtgactgaag ggaagtgtgc cccattcttt 960  
tacggcggat gtggcggcaa ccggaacaac tttgacacag aagagtactg catggccgtg 1020  
tgtggcagcg ccattccctac aacagcagcc agtaccctg atgccgttga caagtatctc 1080  
gagacacctg gggatgagaa tgaacatgcc catttccaga aagccaaaga gaggcttgag 1140  
gccaagcacc gagagagaat gtcccaggte atgagagaat gggaagaggc agaactgag 1200  
gcaaagaact tgcctaaagc tgataagaag gcagttatcc agcatttcca ggagaaagtg 1260  
gaatcttttg aacaggaagc agccaacgag agacagcagc tgggtggagac acacatggcc 1320  
agagtgggaag ccattgctcaa tgaccgccgc cgcctggccc tggagaacta catcaccgct 1380  
ctgcaggctg ttctctctcg gcctcgtcac gtgttcaata tgctaaagaa gtatgtccgc 1440  
gcagaacaga aggacagaca gcacacccta aagcatttgc agcatgtgcg catggtggat 1500  
cccaagaaag ccgctcagat ccggtcccag gttatgacac acctccgtgt gatttatgag 1560  
cgcatgaatc agtctctctc cctgctctac aacgtgcctg cagtggccga ggagattcag 1620  
gatgaagttg atgagctgct tcagaaagag caaaactatt cagatgacgt cttggccaac 1680  
atgattagtg aaccaaggat cagttacgga aacgatgctc tcatgccatc tttgaccgaa 1740  
acgaaaacca ccgtggagct ccttcccgctg aatggagagt tcagcctgga cgatctccag 1800  
ccgtggcatt cttttggggc tgactctgtg ccagccaaca cagaaaacga agttgagcct 1860  
gttgatgccc gccctgctgc cgaccgagga ctgaccactc gaccaggttc tgggttgaca 1920  
aatatcaaga cggaggagat ctctgaagtg aagatggatg cagaattccg acatgactca 1980  
ggatatgaag ttcacatca aaaattggtg ttctttgcag aagatgtggg ttcaaacaaa 2040  
ggtgcaatca ttggactcat ggtgggcggt gttgtcatag cgacagtgat cgtcatcacc 2100  
ttggtgatgc tgaagaagaa acagtacaca tccattcatc atggtgtggt ggaggttgac 2160  
gccgctgtca cccagagga gcgccacctg tccaagatgc agcagaacgg ctacgaaaat 2220  
ccaacctaca agttctttga gcagatgcag aacaagaag 2259

<210> 61  
<211> 753  
<212> PRT  
<213> Homo sapiens

<400> 61  
Met Leu Pro Gly Leu Ala Leu Leu Leu Ala Ala Trp Thr Ala Arg  
1 5 10 15  
Ala Leu Glu Val Pro Thr Asp Gly Asn Ala Gly Leu Leu Ala Glu Pro  
20 25 30

Gln	Ile	Ala	Met	Phe	Cys	Gly	Arg	Leu	Asn	Met	His	Met	Asn	Val	Gln	35	40	45
Asn	Gly	Lys	Trp	Asp	Ser	Asp	Pro	Ser	Gly	Thr	Lys	Thr	Cys	Ile	Asp	50	55	60
Thr	Lys	Glu	Gly	Ile	Leu	Gln	Tyr	Cys	Gln	Glu	Val	Tyr	Pro	Glu	Leu	65	70	75
Gln	Ile	Thr	Asn	Val	Val	Glu	Ala	Asn	Gln	Pro	Val	Thr	Ile	Gln	Asn	85	90	95
Trp	Cys	Lys	Arg	Gly	Arg	Lys	Gln	Cys	Lys	Thr	His	Pro	His	Phe	Val	100	105	110
Ile	Pro	Tyr	Arg	Cys	Leu	Val	Gly	Glu	Phe	Val	Ser	Asp	Ala	Leu	Leu	115	120	125
Val	Pro	Asp	Lys	Cys	Lys	Phe	Leu	His	Gln	Glu	Arg	Met	Asp	Val	Cys	130	135	140
Glu	Thr	His	Leu	His	Trp	His	Thr	Val	Ala	Lys	Glu	Thr	Cys	Ser	Glu	145	150	155
Lys	Ser	Thr	Asn	Leu	His	Asp	Tyr	Gly	Met	Leu	Leu	Pro	Cys	Gly	Ile	165	170	175
Asp	Lys	Phe	Arg	Gly	Val	Glu	Phe	Val	Cys	Cys	Pro	Leu	Ala	Glu	Glu	180	185	190
Ser	Asp	Asn	Val	Asp	Ser	Ala	Asp	Ala	Glu	Glu	Asp	Asp	Ser	Asp	Val	195	200	205
Trp	Trp	Gly	Gly	Ala	Asp	Thr	Asp	Tyr	Ala	Asp	Gly	Ser	Glu	Asp	Lys	210	215	220
Val	Val	Glu	Val	Ala	Glu	Glu	Glu	Glu	Val	Ala	Glu	Val	Glu	Glu	Glu	225	230	235
Glu	Ala	Asp	Asp	Asp	Glu	Asp	Asp	Glu	Asp	Gly	Asp	Glu	Val	Glu	Glu	245	250	255
Glu	Ala	Glu	Glu	Pro	Tyr	Glu	Glu	Ala	Thr	Glu	Arg	Thr	Thr	Ser	Ile	260	265	270
Ala	Thr	Thr	Thr	Thr	Thr	Thr	Thr	Glu	Ser	Val	Glu	Glu	Val	Val	Arg	275	280	285
Glu	Val	Cys	Ser	Glu	Gln	Ala	Glu	Thr	Gly	Pro	Cys	Arg	Ala	Met	Ile	290	295	300
Ser	Arg	Trp	Tyr	Phe	Asp	Val	Thr	Glu	Gly	Lys	Cys	Ala	Pro	Phe	Phe	305	310	315
Tyr	Gly	Gly	Cys	Gly	Gly	Asn	Arg	Asn	Asn	Phe	Asp	Thr	Glu	Glu	Tyr	325	330	335
Cys	Met	Ala	Val	Cys	Gly	Ser	Ala	Ile	Pro	Thr	Thr	Ala	Ala	Ser	Thr	340	345	350



Pro Asp Ala Val Asp Lys Tyr Leu Glu Thr Pro Gly Asp Glu Asn Glu  
 355 360 365  
 His Ala His Phe Gln Lys Ala Lys Glu Arg Leu Glu Ala Lys His Arg  
 370 375 380  
 Glu Arg Met Ser Gln Val Met Arg Glu Trp Glu Glu Ala Glu Arg Gln  
 385 390 395 400  
 Ala Lys Asn Leu Pro Lys Ala Asp Lys Lys Ala Val Ile Gln His Phe  
 405 410 415  
 Gln Glu Lys Val Glu Ser Leu Glu Gln Glu Ala Ala Asn Glu Arg Gln  
 420 425 430  
 Gln Leu Val Glu Thr His Met Ala Arg Val Glu Ala Met Leu Asn Asp  
 435 440 445  
 Arg Arg Arg Leu Ala Leu Glu Asn Tyr Ile Thr Ala Leu Gln Ala Val  
 450 455 460  
 Pro Pro Arg Pro Arg His Val Phe Asn Met Leu Lys Lys Tyr Val Arg  
 465 470 475 480  
 Ala Glu Gln Lys Asp Arg Gln His Thr Leu Lys His Phe Glu His Val  
 485 490 495  
 Arg Met Val Asp Pro Lys Lys Ala Ala Gln Ile Arg Ser Gln Val Met  
 500 505 510  
 Thr His Leu Arg Val Ile Tyr Glu Arg Met Asn Gln Ser Leu Ser Leu  
 515 520 525  
 Leu Tyr Asn Val Pro Ala Val Ala Glu Glu Ile Gln Asp Glu Val Asp  
 530 535 540  
 Glu Leu Leu Gln Lys Glu Gln Asn Tyr Ser Asp Asp Val Leu Ala Asn  
 545 550 555 560  
 Met Ile Ser Glu Pro Arg Ile Ser Tyr Gly Asn Asp Ala Leu Met Pro  
 565 570 575  
 Ser Leu Thr Glu Thr Lys Thr Thr Val Glu Leu Leu Pro Val Asn Gly  
 580 585 590  
 Glu Phe Ser Leu Asp Asp Leu Gln Pro Trp His Ser Phe Gly Ala Asp  
 595 600 605  
 Ser Val Pro Ala Asn Thr Glu Asn Glu Val Glu Pro Val Asp Ala Arg  
 610 615 620  
 Pro Ala Ala Asp Arg Gly Leu Thr Thr Arg Pro Gly Ser Gly Leu Thr  
 625 630 635 640  
 Asn Ile Lys Thr Glu Glu Ile Ser Glu Val Lys Met Asp Ala Glu Phe  
 645 650 655  
 Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe  
 660 665 670  
 Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val

675	680	685
Gly Gly Val Val Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu		
690	695	700
Lys Lys Lys Gln Tyr Thr Ser Ile His His Gly Val Val Glu Val Asp		
705	710	715 720
Ala Ala Val Thr Pro Glu Glu Arg His Leu Ser Lys Met Gln Gln Asn		
725	730	735
Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln Asn Lys		
740	745	750

Lys

<210> 62  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: synthetic

<400> 62  
 Leu Glu Val Leu Phe Gln Gly Pro  
 1 5

<210> 63  
 <211> 10  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: synthetic

<400> 63  
 Ser Glu Val Asn Leu Asp Ala Glu Phe Arg  
 1 5 10

<210> 64  
 <211> 10  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: synthetic

<400> 64  
 Ser Glu Val Lys Met Asp Ala Glu Phe Arg  
 1 5 10

<210> 65  
 <211> 15  
 <212> PRT  
 <213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: synthetic

&lt;400&gt; 65

Arg Arg Gly Gly Val Val Ile Ala Thr Val Ile Val Gly Glu Arg  
 1 5 10 15

&lt;210&gt; 66

&lt;211&gt; 518

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 66

Met Gly Ala Leu Ala Arg Ala Leu Leu Leu Pro Leu Leu Ala Gln Trp  
 1 5 10 15

Leu Leu Arg Ala Ala Pro Glu Leu Ala Pro Ala Pro Phe Thr Leu Pro  
 20 25 30

Leu Arg Val Ala Ala Ala Thr Asn Arg Val Val Ala Pro Thr Pro Gly  
 35 40 45

Pro Gly Thr Pro Ala Glu Arg His Ala Asp Gly Leu Ala Leu Ala Leu  
 50 55 60

Glu Pro Ala Leu Ala Ser Pro Ala Gly Ala Ala Asn Phe Leu Ala Met  
 65 70 75 80

Val Asp Asn Leu Gln Gly Asp Ser Gly Arg Gly Tyr Tyr Leu Glu Met  
 85 90 95

Leu Ile Gly Thr Pro Pro Gln Lys Leu Gln Ile Leu Val Asp Thr Gly  
 100 105 110

Ser Ser Asn Phe Ala Val Ala Gly Thr Pro His Ser Tyr Ile Asp Thr  
 115 120 125

Tyr Phe Asp Thr Glu Arg Ser Ser Thr Tyr Arg Ser Lys Gly Phe Asp  
 130 135 140

Val Thr Val Lys Tyr Thr Gln Gly Ser Trp Thr Gly Phe Val Gly Glu  
 145 150 155 160

Asp Leu Val Thr Ile Pro Lys Gly Phe Asn Thr Ser Phe Leu Val Asn  
 165 170 175

Ile Ala Thr Ile Phe Glu Ser Glu Asn Phe Phe Leu Pro Gly Ile Lys  
 180 185 190

Trp Asn Gly Ile Leu Gly Leu Ala Tyr Ala Thr Leu Ala Lys Pro Ser  
 195 200 205

Ser Ser Leu Glu Thr Phe Phe Asp Ser Leu Val Thr Gln Ala Asn Ile  
 210 215 220

Pro Asn Val Phe Ser Met Gln Met Cys Gly Ala Gly Leu Pro Val Ala  
 225 230 235 240

Gly Ser Gly Thr Asn Gly Gly Ser Leu Val Leu Gly Gly Ile Glu Pro  
 245 250 255

Ser Leu Tyr Lys Gly Asp Ile Trp Tyr Thr Pro Ile Lys Glu Glu Trp  
 260 265 270  
 Tyr Tyr Gln Ile Glu Ile Leu Lys Leu Glu Ile Gly Gly Gln Ser Leu  
 275 280 285  
 Asn Leu Asp Cys Arg Glu Tyr Asn Ala Asp Lys Ala Ile Val Asp Ser  
 290 295 300  
 Gly Thr Thr Leu Leu Arg Leu Pro Gln Lys Val Phe Asp Ala Val Val  
 305 310 315 320  
 Glu Ala Val Ala Arg Ala Ser Leu Ile Pro Glu Phe Ser Asp Gly Phe  
 325 330 335  
 Trp Thr Gly Ser Gln Leu Ala Cys Trp Thr Asn Ser Glu Thr Pro Trp  
 340 345 350  
 Ser Tyr Phe Pro Lys Ile Ser Ile Tyr Leu Arg Asp Glu Asn Ser Ser  
 355 360 365  
 Arg Ser Phe Arg Ile Thr Ile Leu Pro Gln Leu Tyr Ile Gln Pro Met  
 370 375 380  
 Met Gly Ala Gly Leu Asn Tyr Glu Cys Tyr Arg Phe Gly Ile Ser Pro  
 385 390 395 400  
 Ser Thr Asn Ala Leu Val Ile Gly Ala Thr Val Met Glu Gly Phe Tyr  
 405 410 415  
 Val Ile Phe Asp Arg Ala Gln Lys Arg Val Gly Phe Ala Ala Ser Pro  
 420 425 430  
 Cys Ala Glu Ile Ala Gly Ala Ala Val Ser Glu Ile Ser Gly Pro Phe  
 435 440 445  
 Ser Thr Glu Asp Val Ala Ser Asn Cys Val Pro Ala Gln Ser Leu Ser  
 450 455 460  
 Glu Pro Ile Leu Trp Ile Val Ser Tyr Ala Leu Met Ser Val Cys Gly  
 465 470 475 480  
 Ala Ile Leu Leu Val Leu Ile Val Leu Leu Leu Leu Pro Phe Arg Cys  
 485 490 495  
 Gln Arg Arg Pro Arg Asp Pro Glu Val Val Asn Asp Glu Ser Ser Leu  
 500 505 510  
 Val Arg His Arg Trp Lys  
 515

<210> 67  
 <211> 475  
 <212> PRT  
 <213> Homo sapiens

<400> 67  
 Met Gly Ala Leu Ala Arg Ala Leu Leu Leu Pro Leu Leu Ala Gln Trp  
 1 5 10 15  
 Leu Leu Arg Ala Ala Pro Glu Leu Ala Pro Ala Pro Phe Thr Leu Pro

20					25					30						
Leu	Arg	Val	Ala	Ala	Ala	Thr	Asn	Arg	Val	Val	Ala	Pro	Thr	Pro	Gly	
35					40					45						
Pro	Gly	Thr	Pro	Ala	Glu	Arg	His	Ala	Asp	Gly	Leu	Ala	Leu	Ala	Leu	
50					55					60						
Glu	Pro	Ala	Leu	Ala	Ser	Pro	Ala	Gly	Ala	Ala	Asn	Phe	Leu	Ala	Met	
65					70					75					80	
Val	Asp	Asn	Leu	Gln	Gly	Asp	Ser	Gly	Arg	Gly	Tyr	Tyr	Leu	Glu	Met	
85					90					95						
Leu	Ile	Gly	Thr	Pro	Pro	Gln	Lys	Leu	Gln	Ile	Leu	Val	Asp	Thr	Gly	
100					105					110						
Ser	Ser	Asn	Phe	Ala	Val	Ala	Gly	Thr	Pro	His	Ser	Tyr	Ile	Asp	Thr	
115					120					125						
Tyr	Phe	Asp	Thr	Glu	Arg	Ser	Ser	Thr	Tyr	Arg	Ser	Lys	Gly	Phe	Asp	
130					135					140						
Val	Thr	Val	Lys	Tyr	Thr	Gln	Gly	Ser	Trp	Thr	Gly	Phe	Val	Gly	Glu	
145					150					155					160	
Asp	Leu	Val	Thr	Ile	Pro	Lys	Gly	Phe	Asn	Thr	Ser	Phe	Leu	Val	Asn	
165					170					175						
Ile	Ala	Thr	Ile	Phe	Glu	Ser	Glu	Asn	Phe	Phe	Leu	Pro	Gly	Ile	Lys	
180					185					190						
Trp	Asn	Gly	Ile	Leu	Gly	Leu	Ala	Tyr	Ala	Thr	Leu	Ala	Lys	Pro	Ser	
195					200					205						
Ser	Ser	Leu	Glu	Thr	Phe	Phe	Asp	Ser	Leu	Val	Thr	Gln	Ala	Asn	Ile	
210					215					220						
Pro	Asn	Val	Phe	Ser	Met	Gln	Met	Cys	Gly	Ala	Gly	Leu	Pro	Val	Ala	
225					230					235					240	
Gly	Ser	Gly	Thr	Asn	Gly	Gly	Ser	Leu	Val	Leu	Gly	Gly	Ile	Glu	Pro	
245					250					255						
Ser	Leu	Tyr	Lys	Gly	Asp	Ile	Trp	Tyr	Thr	Pro	Ile	Lys	Glu	Glu	Trp	
260					265					270						
Tyr	Tyr	Gln	Ile	Glu	Ile	Leu	Lys	Leu	Glu	Ile	Gly	Gly	Gln	Ser	Leu	
275					280					285						
Asn	Leu	Asp	Cys	Arg	Glu	Tyr	Asn	Ala	Asp	Lys	Ala	Ile	Val	Asp	Ser	
290					295					300						
Gly	Thr	Thr	Leu	Leu	Arg	Leu	Pro	Gln	Lys	Val	Phe	Asp	Ala	Val	Val	
305					310					315					320	
Glu	Ala	Val	Ala	Arg	Ala	Ser	Leu	Ile	Pro	Glu	Phe	Ser	Asp	Gly	Phe	
325					330					335						
Trp	Thr	Gly	Ser	Gln	Leu	Ala	Cys	Trp	Thr	Asn	Ser	Glu	Thr	Pro	Trp	
340					345					350						

Ser Tyr Phe Pro Lys Ile Ser Ile Tyr Leu Arg Asp Glu Asn Ser Ser  
355 360 365

Arg Ser Phe Arg Ile Thr Ile Leu Pro Gln Leu Tyr Ile Gln Pro Met  
370 375 380

Met Gly Ala Gly Leu Asn Tyr Glu Cys Tyr Arg Phe Gly Ile Ser Pro  
385 390 395 400

Ser Thr Asn Ala Leu Val Ile Gly Ala Thr Val Met Glu Gly Phe Tyr  
405 410 415

Val Ile Phe Asp Arg Ala Gln Lys Arg Val Gly Phe Ala Ala Ser Pro  
420 425 430

Cys Ala Glu Ile Ala Gly Ala Ala Val Ser Glu Ile Ser Gly Pro Phe  
435 440 445

Ser Thr Glu Asp Val Ala Ser Asn Cys Val Pro Ala Gln Ser Leu Ser  
450 455 460

Glu Pro Ile Leu Trp His His His His His His  
465 470 475

<210> 68

<211> 413

<212> PRT

<213> Homo sapiens

<400> 68

Ala Leu Glu Pro Ala Leu Ala Ser Pro Ala Gly Ala Ala Asn Phe Leu  
1 5 10 15

Ala Met Val Asp Asn Leu Gln Gly Asp Ser Gly Arg Gly Tyr Tyr Leu  
20 25 30

Glu Met Leu Ile Gly Thr Pro Pro Gln Lys Leu Gln Ile Leu Val Asp  
35 40 45

Thr Gly Ser Ser Asn Phe Ala Val Ala Gly Thr Pro His Ser Tyr Ile  
50 55 60

Asp Thr Tyr Phe Asp Thr Glu Arg Ser Ser Thr Tyr Arg Ser Lys Gly  
65 70 75 80

Phe Asp Val Thr Val Lys Tyr Thr Gln Gly Ser Trp Thr Gly Phe Val  
85 90 95

Gly Glu Asp Leu Val Thr Ile Pro Lys Gly Phe Asn Thr Ser Phe Leu  
100 105 110

Val Asn Ile Ala Thr Ile Phe Glu Ser Glu Asn Phe Phe Leu Pro Gly  
115 120 125

Ile Lys Trp Asn Gly Ile Leu Gly Leu Ala Tyr Ala Thr Leu Ala Lys  
130 135 140

Pro Ser Ser Ser Leu Glu Thr Phe Phe Asp Ser Leu Val Thr Gln Ala  
145 150 155 160

Asn Ile Pro Asn Val Phe Ser Met Gln Met Cys Gly Ala Gly Leu Pro

```

<210> 69
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Peptide

<400> 69
Gly Leu Ala Leu Ala Leu Glu Pro
  1               5

```

<211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Peptide

<400> 70  
 Glu Val Lys Met Asp Ala Glu Phe  
           1                          5

<210> 71  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Peptide

<400> 71  
 Glu Val Asn Leu Asp Ala Glu Phe  
           1                          5

<210> 72  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Peptide

<400> 72  
 Leu Val Phe Phe Ala Glu Asp Val  
           1                          5

<210> 73  
 <211> 8  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Peptide

<400> 73  
 Lys Leu Val Phe Phe Ala Glu Asp  
           1                          5

<210> 74  
 <211> 38  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer

<400> 74  
 cgctttaagc ttgccaccat gggcgactg gcccgggcg

<210> 75



<211> 57  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 75  
cgctttctcg agctaattgt gatggtgatg gtgccacaaa atgggctcgc tcaaaga 57

<210> 76  
<211> 15  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Peptide

<400> 76  
Arg Arg Gly Gly Val Val Ile Ala Thr Val Ile Val Gly Glu Arg  
1 5 10 15

What is claimed is:

1. A method for assaying hu-Asp1  $\alpha$ -secretase activity comprising the steps of:
  - (a) contacting hu-Asp1 protein with an amyloid precursor protein (APP) substrate, wherein said substrate contains an  $\alpha$ -secretase cleavage site; and
  - (b) measuring cleavage of the APP substrate at the  $\alpha$ -cleavage site, thereby assaying hu-Asp1  $\alpha$ -secretase activity.
2. A method according to claim 1, wherein the hu-Asp1 protein comprises a polypeptide produced in cell transformed or transfected with a polynucleotide comprising a nucleotide sequence that encodes hu-Asp1 or a fragment thereof that retains Asp1  $\alpha$ -secretase activity.
3. A method of claim 2, wherein the hu-Asp1 protein is purified and isolated from said cell.
4. A method according to claims 2 or 3, wherein the nucleotide sequence encodes a polypeptide that comprises the hu-Asp1 amino acid sequence set forth in SEQ ID NO: 2 or a fragment thereof, wherein said fragment retains  $\alpha$ -secretase activity.
5. A method according to any one of claims 2-4, wherein the polynucleotide sequence encodes a hu-Asp1 amino acid sequence lacking the transmembrane amino acids 469-492 of SEQ ID NO: 2.
6. A method according to claim 5, wherein the polynucleotide sequence encodes a hu-Asp1 amino acid sequence further lacking the cytoplasmic domain amino acids 493-518 of SEQ ID NO: 2.

7. A method according to any one of claims 1-6, wherein the hu-Asp1 amino acid sequence lacks amino terminal amino acids 1-62 of SEQ ID NO: 2.
8. A method according to claims 1 or 2, wherein the contacting step comprises growing a cell transfected or transformed with a polynucleotide encoding hu-Asp1 protein or a fragment thereof that retains hu-Asp1  $\alpha$ -secretase activity, wherein the cell is grown under conditions in which the cell expresses the hu-Asp1 protein in the presence of the APP substrate.
9. A method of claim 8, wherein said cell expresses a polynucleotide encoding an APP substrate containing an  $\alpha$ -secretase cleavage site, and wherein the contacting step comprises growing the cell under conditions in which the cell expresses the hu-Asp1 protein and the APP substrate.
10. A method of any one of claims 1-9, wherein the APP substrate  $\alpha$ -secretase cleavage site comprises the amino acid sequence LVFFAEDF or KLVFFAED.
11. A method of any one of claims 1-10, wherein the APP substrate comprises a human APP isoform and further comprises a carboxy-terminal di-lysine.
12. A method of claims 10 or 11, wherein the APP substrate comprises a detectable label.
13. A method of claim 12, wherein the detectable label is selected from the group consisting of radioactive labels, chemiluminescent labels, enzymatic labels, chemiluminescent labels and fluorescent labels.

14. A method of any one of claims 1-13, wherein the APP substrate comprises a human APP isoform and the determining step comprises measuring the production of amyloid alpha peptide (sAPP $\alpha$ ).

15. A method of any one of claims 1-14, wherein the method further comprises the steps of:

- (a) determining the level of hu-Asp1  $\alpha$ -secretase activity in the presence and absence of a modulator of hu-Asp1  $\alpha$ -secretase activity; and
- (b) comparing the hu-Asp1  $\alpha$ -secretase activity in the presence and absence of the modulator, wherein modulators that increase hu-Asp1  $\alpha$ -secretase activity are identified as candidate Alzheimer's disease therapeutics.

16. A method of claim 15, wherein the method further comprises a step of treating Alzheimer's disease with said candidate Alzheimer's disease therapeutic.

17. A composition comprising a candidate Alzheimer's disease therapeutic identified by the method of claim 15.

18. A hu-Asp1 protease substrate peptide or fragment thereof, wherein said peptide comprises an amino acid sequence consisting of fifty or fewer amino acids, said amino acid sequence including the hu-Asp1 cleavage site having the amino acid sequence GLALALEP

19. The substrate of claim 18, wherein the substrate comprises a detectable label.

20. The substrate of claim 19, wherein the detectable label is selected from the group consisting of radioactive labels, enzymatic labels, chemiluminescent labels and fluorescent labels.

21. A method for assaying hu-Asp1 proteolytic activity comprising the steps of:
- (a) contacting hu-Asp1 protein with an hu-Asp1 substrate according to claim 18, 19, or 20 under acidic conditions, and
  - (b) determining the level of hu-Asp1 proteolytic activity.
22. A method according to claim 20, wherein the hu-Asp1 protein comprises a polypeptide produced in cell transformed or transfected with a polynucleotide comprising the nucleotide sequence that encodes hu-Asp1.
23. A method of claim 21, wherein the hu-Asp1 protein is purified and isolated from said cell.
24. A method according to claims 22 and 23, wherein the nucleotide sequence encodes a polypeptide that comprises the hu-Asp1 amino acid sequence set forth in SEQ ID NO: 2 or a fragment thereof, wherein said fragment retains proteolytic activity.
25. A purified polynucleotide comprising a nucleotide sequence encoding a polypeptide that comprises a fragment of a hu-Asp-1 protein, wherein said nucleotide sequence lacks the sequence that encodes the transmembrane domain of said hu-Asp1 protein, and wherein the hu-Asp1 polypeptide fragment encoded by said polynucleotide has hu-Asp1  $\alpha$ -secretase activity.
26. A polynucleotide of claim 25, wherein the polypeptide comprises a fragment of the hu-Asp1 amino acid sequence set forth in SEQ ID NO: 2, and wherein the polypeptide lacks the transmembrane domain amino acids 469-492 of SEQ ID NO: 2.

27. A polynucleotide of claim 26 wherein the polypeptide further lacks cytoplasmic domain amino acids 493-518 of SEQ ID NO: 2.

28. A purified polynucleotide comprising a nucleotide sequence encoding a polypeptide that comprises a fragment of a hu-Asp1 protein, wherein said nucleotide sequence lacks the sequence that encodes amino acids 1-62 of SEQ ID NO: 2, and wherein the polypeptide has hu-Asp1  $\alpha$ -secretase activity.

29. A polypeptide of claim 28, wherein said polypeptide lacks amino terminal amino acids 23-62 of SEQ ID NO: 2.

30. A purified polynucleotide comprising a nucleotide sequence that hybridizes under stringent conditions to the non-coding strand complementary to SEQ ID NO: 1, wherein the nucleotide sequence encodes a polypeptide having Asp1 proteolytic activity and wherein the polynucleotide lacks nucleotides encoding a transmembrane domain.

31. A purified polynucleotide comprising a nucleotide sequence that hybridizes under stringent conditions to the nucleic acid according to claim 30, wherein the nucleotide sequence encodes a polypeptide further lacking a pro-peptide domain corresponding to amino acids 23-62 of SEQ ID NO: 2.

32. A vector comprising the polynucleotide of any one of claims 25-31.

33. A host cell transformed or transfected with a vector of claim 32.

34. A host cell transformed or transfected with a polynucleotide of any one of claims 25-31.

35. A purified polypeptide comprising a fragment of a hu-Asp1 protein, wherein said polypeptide lacks the hu-Asp1 transmembrane domain of said hu-Asp1 protein and retains hu-Asp1  $\alpha$ -secretase activity.

36. A polypeptide of claim 35, wherein said polypeptide comprises a fragment of hu-Asp1 having the amino acid sequence set forth in SEQ ID NO: 2 and wherein said polypeptide lacks the transmembrane domain amino acids 469-492 of SEQ ID NO: 2.

37. A polypeptide according to claim 35 or 36, wherein said polypeptide further lacks cytoplasmic domain amino acids 493-518 of SEQ ID NO: 2.

38. A polypeptide according to any one of claims 35-37, which further lacks amino terminal amino acids 1-62 of SEQ ID NO: 2.

39. A polypeptide comprising a fragment of hu-Asp1 having the amino acid sequence set forth in SEQ ID NO: 2 and wherein said polypeptide lacks the amino terminal amino acids 1-62 of SEQ ID NO: 2 and retains APP processing activity.

40. A polypeptide comprising an amino acid sequence at least 95% identical to a fragment of hu-Asp1 protein, wherein said polypeptide and said fragment lack a transmembrane domain and retain hu-Asp1  $\alpha$ -secretase activity.

41. A polypeptide comprising an amino acid sequence at least 95% identical to a fragment of hu-Asp1 protein, wherein said polypeptide and said fragment lack the amino terminal amino acids corresponding to the pro-peptide domain of hu-Asp1 and retain APP processing activity.



168



INVESTOR IN PEOPLE

Application No: GB 0023315.5 Examiner: Dr Patrick Purcell  
Claims searched: 1-15, 17, 25-29, 35-41 and in Date of search: 27 April 2001  
part 32-34

## Patents Act 1977 Search Report under Section 17

### Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.S):

Int Cl (Ed.7):

Other: ONLINE: EPODOC, WPI, JAPIO, DGENE, CAPLUS, BIOSIS, MEDLINE

### Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	EP 0848062 A2 (SMITHKLINE BEECHAM CORP. & SMITHKLINE BEECHAM PLC) see whole document, esp. page 2, lines 25-30 and Tables 1, 2	
E, A	DE 19910108 A1 (FAHRENHOLZ) see whole document and WPI abstract no. 2000-588391	
E, A	DE 19849073 A1 (AVENTIS PHARMA DEUTSCHLAND GMBH) see whole document, esp. page 4, lines 4-9 and WPI abstract No. 2000-304753	
A	J. Neurochem., vol 71, 1998, GK Gouras et al, "Generation and regulation of $\beta$ -amyloid peptide variants by neurons", 1920-1925, esp. Fig. 1B	

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.